A STATISTICAL ANALYSIS OF FACTORS ASSOCIATED WITH DRIVER-PERCEIVED ROAD ROUGHNESS ON URBAN HIGHWAYS

WA-RD 538.1

Research Report
June 2002

Washington State Department of Transportation

Washington State Transportation Commission
Planning and Capital Program Management
in cooperation with
U.S. DOT – Federal Highway Administration

Research Report
Research Project T1803, Task 28
Highway Roughness

A STATISTICAL ANALYSIS OF FACTORS ASSOCIATED WITH DRIVER-PERCEIVED ROAD ROUGHNESS ON URBAN HIGHWAYS

By

Kevan Shafizadeh
Doctoral Research Assistant
Civil and Environmental Engineering
University of Washington
Seattle, Washington 98195

Fred Mannering
Professor
Civil and Environmental Engineering
Purdue University
West Lafayette, Indiana 47907

Linda Pierce
Pavement and Soils Engineer
Washington State Department of Transportation

Washington State Transportation Center (TRAC)
University of Washington, Box 354802
University District Building
1107 NE 45th Street, Suite 535
Seattle, Washington 98105-4631

Washington State Department of Transportation
Technical Monitor
Linda Pierce
Pavement and Soils Engineer

Prepared for

Washington State Transportation Commission
Department of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

June 2002

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
WA-RD 538.1		

4. TITLE AND SUBTITLE	5. REPORT DATE
A STATISTICAL ANALYSIS OF FACTORS ASSOCIATED WITH DRIVER-PERCEIVED ROAD ROUGHNESS ON URBAN HIGHWAYS	June 2002
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Kevan Shafizadeh, Fred Mannering, Linda Pierce	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NO.
Washington State Transportation Center (TRAC) University of Washington, Box 354802 University District Building; 1107 NE 45th Street, Suite 535 Seattle, Washington 98105-4631	11. CONTRACT OR GRANT NO. Agreement T1803, Task 28
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Research Office Washington State Department of Transportation Transportation Building, MS 47370 Olympia, Washington 98504-7370 Keith Anderson, Project Manager, 360-709-5405	Final Research Report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

This study linked physical quantitative roadway measures with subjective, driver-perceived measures of roughness on urban highways. 56 participants from the general public were placed in normal traffic conditions and asked their opinions about pavement roughness on 40 predetermined highway test segments. Driver evaluations were collected with other data, such as speed and in-vehicle noise, and matched with driver-specific socio-demographic data and pavement-specific data from the Washington State Department of Transportation and its pavement management system.

Results from an ordered logit model indicated that the international roughness index (IRI) is the single best predictor of driver-perceived road roughness and driver acceptability. Pavements with low IRI values generally corresponded with low roughness rankings and high levels of user acceptability. Other factors statistically associated with driver-perceived measures of road roughness included measured IRI, the presence of pavement maintenance, the presence of joints or bridge abutments, the age of the pavement surface, the vehicle type, levels of in-vehicle noise, the speed of vehicle, and the gender and income of the driver.

This study also provided empirical data that can be used to both support or challenge the federal IRI acceptability threshold of 170 in/mi (2.7 m/km) recommended by the Federal Highway Administration.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Pavement preservation, roughness, international roughness index, ride quality, pavement management systems	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None		

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Transportation Commission, Department of Transportation, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Note: All text, tables, and graphs are presented in U.S. Customary (English) units along with metric (Imperial) units whenever possible. In all other instances, U.S. Customary (English) units are used in the text and duplicate figures using metric units are included in the appendix.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AC – asphaltic concrete, or asphalt concrete pavement (**ACP)**

ANOVA – analysis of variance

AASHTO – American Association of Highway and Transportation Officials, formerly American Association of Highway Officials (**AASHO**)

BPR – Bureau of Public Roads (currently **FHWA**)

BST – bituminous surface treatment

DOT – Department of Transportation

FHWA – Federal Highway Administration

HCM – Highway Capacity Manual

HPMS – Highway Performance Monitoring System

IRI – International Roughness Index

JLARC – Joint Legislative Audit and Review Committee

LOS – level of service, as defined in the **HCM**

MRM - Mays Ride Meter

NCHRP – National Cooperative Highway Research Program

NHS – National Highway System

NHUS – National Highway User Survey

NPHQ – National Partnership for Highway Quality, formerly the federal National Quality Initiative (**NQI**)

PCC – portland cement concrete, or portland cement concrete pavement (**PCCP**)

PMS – pavement management system

PQI – Pavement Quality Index

PPC – Pavement Profile Condition

PRC – Pavement Rutting Condition

PSC – Pavement Structural Condition

PSI – Present Serviceability Index

PSR – present serviceability rating

RTRRMS – response-type road roughness measuring system

SST – Statistical Software Technique software

TRB – Transportation Research Board

TTI – Texas Transportation Institute

WSDOT – Washington State Department of Transportation (**DOT**)

WSPMS – Washington State Pavement Management System (**PMS**)

EXECUTIVE SUMMARY

Understanding and measuring road roughness are key concerns in state and national arenas. Road roughness has been used as a basis for highway resource allocation and forecasting highway needs, and it is also known to affect public satisfaction with the highway system. From a quantitative perspective, the Washington State Department of Transportation (WSDOT) collects detailed information on pavement conditions, including roughness, rutting, faulting, cracking, patching, and scaling.

In Washington State, and nationwide, one extensively used quantifiable measure of roughness is the International Roughness Index (IRI). This index is also used by the Federal Highway Administration (FHWA) to assess changes in the condition of the nation's highways and to forecast highway investment needs. Work on linking the IRI with the motoring public's perception of roughness has been limited. The intent of this study was to quantify such a link.

In this study, drivers were placed in real world driving scenarios and asked to reveal their opinions about pavement roughness. A total of 56 participants each evaluated 40 highway test segments and produced 2,180 separate "observations." Each "observation" contained a driver-perceived roughness ranking and acceptability evaluation, which served as dependent variables, as well as associated quantitative socio-demographic data for the individual from the preliminary survey, physical segment-specific data from the WSPMS (Washington State Pavement Management System), and in-vehicle survey data from the driving experiment, which served as independent variables.

In addition to descriptive statistics, analysis of variance (ANOVA) and ordered probability models were used to identify differences among mean roughness rankings and to identify the relationship between roughness rankings and a set of explanatory variables, respectively. ANOVA techniques were used to identify whether there were significant differences among vehicle type and start location on the basis of speed measures. The ordered probability models linked the roughness rankings to explanatory variables related to the vehicle type, as well as to roadway and driver characteristics.

This study identified some of the factors that were associated with changes in perceived roughness. By using a combination of individual data, pavement data, and revealed roughness rankings, ordinal and binary logit models were estimated. The models probed the factors associated with differences in driver roughness rankings, as well as roadway roughness acceptability. Both models exhibited significant improvements to the log-likelihood function and appeared to yield favorable goodness-of-fit statistics. The findings indicated that, while measured IRI levels provided the strongest indication of rankings and acceptability (as one would expect), other factors associated with the roadway, the test vehicle, and the driver's characteristics were significant. This study identified salient factors associated with drivers' perception of road roughness on urban highways. IRI was shown to be the most significant factor associated with changes in drivers perceptions of road roughness, as well as driver's acceptability of a roadway's condition. All factors are shown in the following table.

Variables Associated with More Roughness	Variables Associated with Less Roughness
• Measured IRI • Observable "maintenance" • Presence of joints/abutments • Age of surface • In-vehicle noise • Vehicle speed • High income users • Male users • Frequent users of I-405	• Older individuals • Sport utility test vehicles • Minivan test vehicles • Female users • Frequent users of SR 520

This study also provided empirical data that can be used to support an IRI acceptability threshold. The IRI guideline recommended by the 1998 *National Strategic Plan* set forth by FHWA is 170 in/mi (2.7 m/km), and this study found that approximately 85 percent of all "acceptable" evaluations fell at or below this value.

One caution concerning the results of this study should be noted. The characteristics of the drivers, as well as the characteristics of the test segments, may be too unique to extrapolate these results statewide or nationally.

CHAPTER 1. INTRODUCTION

PROBLEM STATEMENT AND STUDY MOTIVATION

Understanding and measuring road roughness are key concerns in state and national arenas. Road roughness has been used as a basis for highway resource allocation and forecasting highway needs, and it is also known to affect public satisfaction with the highway system. From a quantitative perspective, the Washington State Department of Transportation (WSDOT) collects detailed information on pavement condition,s including roughness, rutting, faulting, cracking, patching, and scaling.

In Washington State, and nationwide, one extensively used quantifiable measure of roughness is the International Roughness Index (IRI). This index is also used by the Federal Highway Administration (FHWA) to assess changes in the condition of the nation's highways and to forecast highway investment needs. Work on linking the IRI with the motoring public's perception of roughness has been limited. The intent of this study was to quantify such a link.

In 1997, the Washington State legislature passed a law (ESSB 6061) requiring that a series of internal audits be performed on the Department of Transportation. The audits, which were conducted for the Joint Legislative Audit and Review Committee (JLARC), led to a 1997 telephone survey of 508 Washington State residents that indicated that "poor road surface" ranked second only to "congestion/inadequate capacity issues" as the state's biggest transportation problem (Elway Research, 1997). The telephone survey also revealed that

> Among five specific highway elements posed to survey respondents, road surface had the highest negative ratings, with 15 percent of respondents

> indicating an inadequate rating. Fourteen percent of respondents claimed that Washington's highways are either not as good as, or much worse than, other states' highways. (p.2-18, Cambridge Systematics, 1998)

The JLARC audit went on to conclude that while WSDOT has a pavement management system (PMS) that "has the analytic capabilities to help prioritize pavement projects, there are issues with how PMS applications and results are communicated" and that there is "a need for greater recognition of customer perceptions of pavement condition" (p. 2-20, Cambridge Systematics, 1998). Finally, the JLARC report made the following recommendation: "The Washington State Department of Transportation should consider including pavement roughness, in addition to Pavement Structural Condition and rutting, in its candidate pavement project thresholds" (p. 2-20, Cambridge Systematics, 1998). This finding is of interest in that WSDOT does measure and consider roughness in its assessment of pavement condition. This study attempted to address some of these issues by linking the driver's perceptions of surface roughness with measures of roughness.

While this study focused on issues in Washington State, it is important to note that any pavement-related findings could have national implications. In 2000, the Federal Highway Administration duplicated a 1995 study by the National Partnership for Highway Quality (NPHQ), formerly known as the National Quality Initiative (NQI), to gauge the public's satisfaction with the nation's highway system. The results indicated that satisfaction with the pavement conditions of the national highway system remains low, as shown in Figure 1 (Keever et al., 2001; Coopers and Lybrand, 1995). There was some improvement from 1995: "The percentage of miles on the National Highway System (NHS) with an acceptable ride quality (based on an International Roughness

Index [IRI] value of less than 170 in/mi) increased from 90.0 percent to 93.0 percent from 1995 to 1999" (p. 6, Keever et al., 2001). However, when asked which highway characteristic should receive the most attention and resources for improvement, respondents chose pavement conditions (21 percent) only behind improvements to traffic flow (28 percent) and safety (26 percent), as shown in Figure 2.

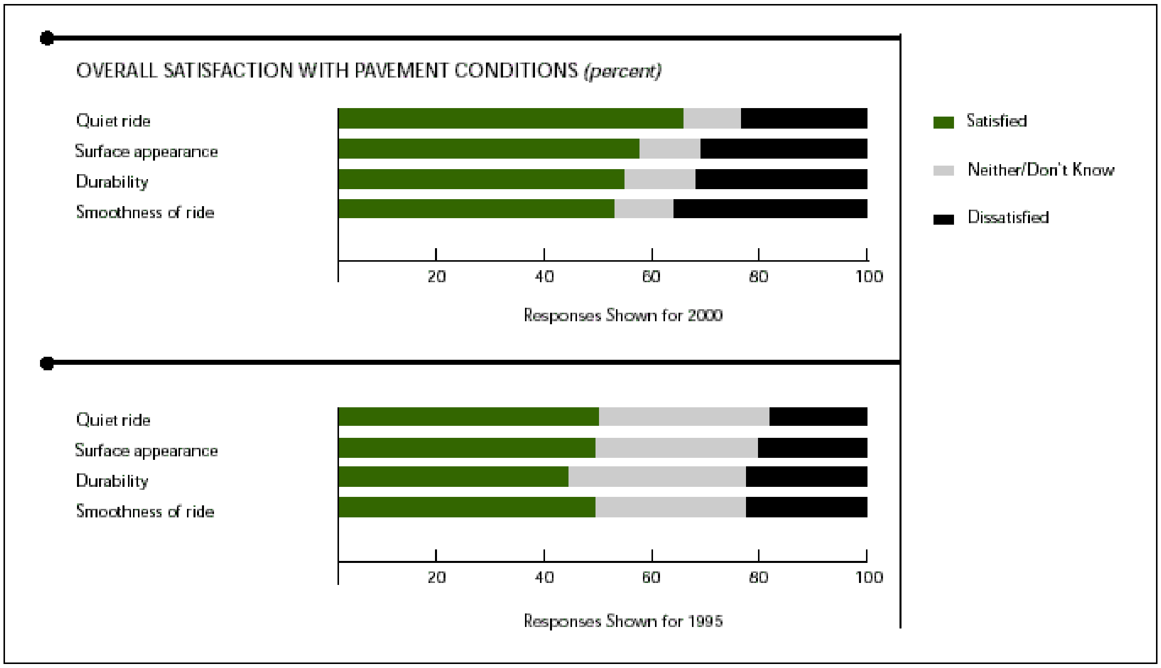


Figure 1. National Satisfaction with Pavement Conditions
Source: Keever et al., 2001.

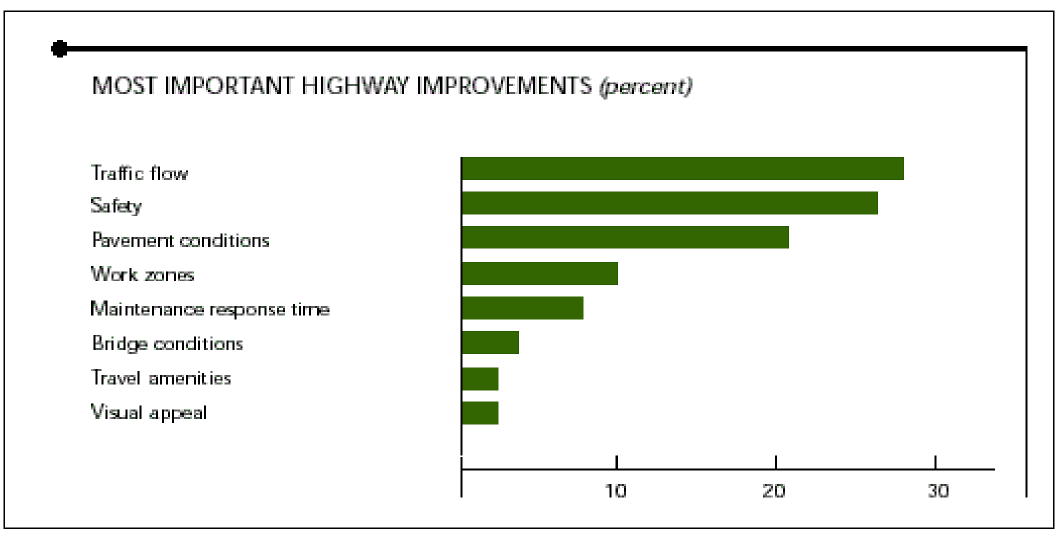


Figure 2. Most Important Highway Improvements According to Users
Source: Keever et al., 2001.

PROJECT SCOPE AND OBJECTIVES

This study had four primary objectives. The first objective was to design an experiment that would link roughness data to public perceptions of road roughness. The second objective was to collect data on the public's general perception of pavement roughness in Washington State. The third was to compare the public's perceptions with actual measurements of road roughness and physical roadway attributes. The last objective was to compare these findings with those in other related research.

PROJECT BENEFITS

The benefits of this study are two-fold: 1) to develop a method of correlating physical measures of roadway roughness with public perceptions of roughness and 2) to better understand what the motoring public defines as "rough" and unacceptable pavement conditions. By understanding the motoring public's needs, WSDOT can better develop strategies to address the public's perceived levels of highway roughness.

Some transportation agencies have been discussing the merit of providing incentives to contractors who can construct a road with a high level of smoothness. One example of this "smoothness" incentive is demonstrated in Maricopa County, Arizona, and is detailed in Table 1. According to Boeger and Crowe, "contractors, under this incentive program, can earn as much as an additional 10 percent of the total project paving costs in incentive bonuses by exceeding the standard preset standard for smoothness" (p. 17, 2002).

Table 1. Road Construction Incentives Based on IRI in Maricopa County, Arizona

IRI (in/mi)	IRI (m/km)	Percent Adjustment
< 51	< 0.80	+10%
51 – 60	0.80 – 0.95	+5%
61 – 80	0.96 – 1.26	0
81 – 101	1.27 – 1.58	-5%
101 – 110	1.59 – 1.74	-10%
111 – 120	1.75 – 1.89	-25%
> 120	> 1.89	Replacement Required

Payment to the contractor is based on the IRI according to this table. The percentage of adjustment will be applied to payments for the total quantity of hot-mix asphalt used in travel lanes only upon completion of the final course of pavement. Source: Boeger and Crowe, 2002.

While this incentive program may provide beneficial quality control, which, in turn, may have an impact on the structural integrity of the road, this research could provide some justification for supporting or questioning the IRI threshold values between incentive levels.

ORGANIZATION OF REPORT

This report is organized into seven chapters. In the second chapter, the reader is presented with a literature review of the research conducted to relate roughness indices and drivers' perceptions of roughness. The third and fourth chapters are devoted to detailing the data collection procedures and subsequent data analysis techniques, respectively. In the fifth chapter, results from the data analyses are presented and findings are developed, before being compared to other research in chapter six. Finally, conclusions and recommendations are made.

CHAPTER 2. BACKGROUND AND LITERATURE REVIEW

This chapter reviews the salient road roughness research in the literature by highlighting the major steps in the evolution of roughness measurement systems and detailing the significant research linking physical roadway measurements with subjective response ratings. The chapter begins by reviewing the early road tests spearheaded by the federal government, then identifies the significant follow-up research that was produced in response to original federal experiments.

THE BEGINNING: THE PAVEMENT SERVICEABILITY CONCEPT

The American Association of State Highway Officials (AASHO) conducted the first research into road roughness in 1960 when it completed the AASHO Road Tests, in which 100 individuals subjectively rated segments of pavements in three states (Illinois, Indiana, and Minnesota) on a scale from 0 to 5 (Carey and Irick, 1960). These ratings were compared to physical measurements obtained from a Bureau of Public Roads (BPR) profilometer.[1] Two models were developed – one for asphaltic concrete (AC) surfaces and one for portland cement concrete (PCC) surfaces – to convert profilometer data into a subjective ratings.

This study was significant because it led to the formulation and development of the concept of pavement "serviceability." The present serviceability rating (PSR), defined by Carey and Irick as "the mean individual ratings made by the members of a specific panel… [who are] intended to represent all highway users," was used to establish

[1] The earliest roughness measurements were reported in 1923, and the BPR roughometer was developed in 1926 (Hudson, 1981). Since then, there have been a handful of roughness measurement instruments, such as roughometers and profilometers.

a present serviceability index (PSI), defined as "a mathematical combination of values obtained from certain physical measurements of a large number of pavements so formulated as to predict the PSR for those pavements within prescribed limits" (p. 42, 1960). In short, the "serviceability" of a highway can be best expressed as the mean roughness rating (i.e., a PSR) given by a panel of passengers in a vehicle, and the PSI was a mathematical model used to convert physical measurements into PSR values.

This new "performance measure" was widely accepted among the transportation (highway) community and became one of the Road Test's most notable contributions. It was among the first successful attempts to standardize roughness ratings to such an extent that they became a reputable performance measure to which actual road roughness measurements were compared. Prior to 1993, all pavement conditions were reported to the FHWA using PSR values (FHWA, 2001).

THE DEVELOPMENT OF THE IRI STANDARD

In the years that followed the original AASHO Road Test, many studies were done to replicate the findings, often using various physical roughness measurement systems (e.g., roughometers, profilometers, ride meters), many of which were not easily comparable at the time because there was a "general lack of equivalence between many methods and measures by which road roughness [was] characterized" (p. 76, Sayers et al., 1986). As noted by Hudson in 1981:

> Diverse measurements of roughness are used around the world. Comparison of equality among measurements is not feasible because no roughness measurement system is capable of giving equal results for all conditions. Rather it is essential that we have compatible measurements.

As a result, some of the general roughness research at the time was devoted to comparing competing roughness measurements:

> Given the multitude of roughness statistics now in use to estimate pavement serviceability, and given the lack of perfect agreement between them, the most practical first step towards meeting the goals of this program is the selection of the best measure for use by everyone and the abandonment of the rest. The commended measure should be the one that best reflects pavement serviceability but at the same time can be adopted by agencies using other measures, with a minimum effort. (p. 32, Gillespie et al., 1980)

Finally in 1982, the World Bank commissioned an experiment in Brazil to establish a roughness measurement standard, and the result was the International Roughness Index (IRI) (Sayers et al., 1986). IRI is now considered the international standard for comparing roughness measurements.[2] In short, the IRI measures the cumulative deviation from a smooth surface in inches per mile (or meters per kilometer).

The International Roughness Index is a mathematical algorithm that takes the profile measurements from a number of "response-type road roughness measuring systems" (RTRRMSs) and calculates the suspension deflection that would be observed from the standard corner suspension of a car (known as a "quarter car"), as shown in Figure 3. The simulated suspension motion is accumulated and divided by the distance traveled to give an index with units of slope, typically m/km or in/mi (Gillespie, 1992).

---

[2] Since 1990, the Federal Highway Administration has required states to report road roughness on the IRI scale, which was later incorporated into the Highway Performance Monitoring System (HPMS).

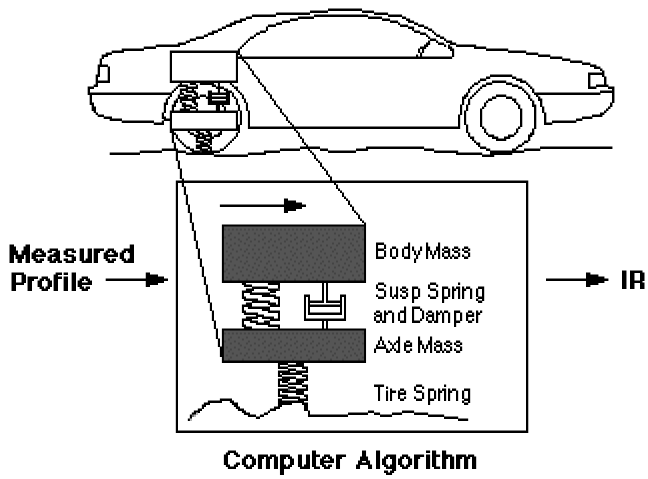


Figure 3. Quarter-Car Model Profile Conversion to IRI

Source: Gillespie, 1992.

Typical IRI values range from 0 to 5 m/km (317 in/mi), with higher values indicating rougher pavement surface. Table 2 contains a qualitative pavement condition term and corresponding quantitative PSR and IRI values.[3] Table 2 also indicates the FHWA descriptive term for pavement condition, "acceptable ride quality," introduced in the 1998 FHWA *National Strategic Plan.* This plan stated that by 2008, 93 percent of the National Highway System (NHS) mileage should meet pavement standards for "acceptable ride quality." In order to be rated "acceptable," pavement performance must have an IRI value of less than or equal to 170 in/mi (2.7 m/km).[4] The term "less than acceptable" is used to describe mileage that does not meet the "acceptable" threshold on

[3] The translation between PSR and IRI is not exact. The IRI values are based on objective measurements of pavement roughness, while PSR is a more subjective evaluation of a broader range of pavement characteristics. For example, a given pavement section could have an IRI value of 165 and be rated as "mediocre," but might be rated a 2.5 and "poor" on the subjective PSR scale.

[4] IRI is required to be reported for all NHS routes, but PSR data are not used to determine "acceptable ride quality" in the *Strategic Plan* or related annual reports. Other principal measures of pavement condition or distress such as rutting, cracking, and faulting are not reported in the Highway Pavement Management System (HPMS) (FHWA, 2001). States vary in the inventories of these distress measures for their highway system, but Washington State is among the national leaders in the development of its pavement management system. To continue to improve pavement evaluation, FHWA, the American Association of Highway and Transportation Officials (AASHTO), and state departments of transportation continue to work to establish standards for measuring roughness, cracking, rutting, and faulting.

the National Highway System. While the threshold of 170 in/mi (2.7 m/km) is clear, it is not clear why this threshold value was set or if there are data to support it. As will be shown, this report fills a major gap in the literature by providing the empirical data to identify driver acceptability (or unacceptability) thresholds.

Table 2. FHWA Pavement Roughness Thresholds for Interstate Facilities

Condition Term	PSR Rating	IRI	NHS Ride Quality
Very Good	≥ 4.0	< 60 in/mi (< 0.95 m/km)	Acceptable: 0 – 170 (in/mi)
Good	3.5 - 3.9	60 – 94 in/mi (0.95 – 1.48 m/km)	
Fair	3.1 - 3.4	95 – 119 in/mi (1.50 – 1.88 m/km)	
Mediocre	2.6 - 3.0	120 – 170 in/mi (1.89 – 2.68 m/km)	
Poor	≤ 2.5	> 170 in/mi (> 2.68 m/km)	Less than Acceptable: > 170 (in/mi)

Source: Federal Highway Administration, 2001.

THE PROBLEM WITH THE SERVICEABILITY CONCEPT

Even with an IRI standard, physical roughness measurement systems remain inadequate to fully model driver perceptions of roughness. As Janoff and Nick pointed out, although Carey and Irick's models exhibited good correlation statistics between subjective ratings and physical measurements, "there are a number of factors that would potentially affect this correlation… [such as] vehicle characteristics,… vehicle operating speed, and composition of the [individual]… [such as] sex, age,…" (p.112, 1983).

Moreover, the same argument can be made about most physical roughness measurements. Currently, the pavement structural condition (PSC), the depth of rutting, and IRI are tracked in the Washington State Pavement Management System (WSPMS), and any criterion can trigger the pavement section for rehabilitation (Sivaneswaran et al., 2000).[5] WSDOT's experience has been that "the PSC or the rut depth will signal the need for rehabilitation before roughness reaches the point of requiring corrective action" (p. 2-14, Cambridge Systematics, 1998). For instance, of the approximately 2870 lane miles of pavements (excluding ramps) determined to need rehabilitation during the 2001-2003 biennium, only 80 lane miles, or less than 3 percent, are determined to need rehabilitation because of roughness.

While the original AASHO Road Tests were designed to capture some of the subjective aspects of road roughness, little research has been done to fill this gap in the literature. The link between physical roughness measurements and individual perception of roughness needs to be better established.

As will eventually be shown, this research extends the existing body of knowledge by linking physical measurements and individual satisfaction: 1) with a respectable set of individual, vehicle, and roadway segment characteristics, 2) through more advanced statistical analysis techniques, and 3) with driver evaluators as subjects.

LITERATURE REVIEW: LINKING PHYSICAL MEASUREMENTS TO PASSENGER PERCEPTIONS OF ROUGHNESS

After the AASHO Road Test, many studies attempted to compare results using slightly modified experimental procedures and/or different or updated road roughness measuring devices in various states (e.g., Nakamura and Michael, 1963; Scrivner and Hudson, 1964; Karan et al., 1979; Moore et al., 1987; Arterbrun and Suprenant, 1990; Ward et al., 1993). These studies produced a number of sets of linear regression

[5] The pavement structural condition (PSC), which monitors the deterioration of pavement, has values ranging from 100 (excellent) to zero (completely deteriorated pavement). The depth of rutting, which consists of channel-shaped depressions in the wheelpaths of the pavement, is measured in millimeters.

equations for predicting panel ratings (PSRs) using some form of roughness measurement. All of these studies are summarized at the end of this chapter in Table 8.

Nakamura and Michael (1963)

The research by Nakamura and Michael at Purdue University in 1963 is a good example of the early research that followed the 1960 Road Test. Sixty test sections (19 rigid, 22 composite, and 19 flexible) located within a 40-mile radius of Lafayette, Indiana, were evaluated by 30 passenger raters. The 30 passenger raters were divided into three panels – one composed of Indiana State highway engineers, one composed of civil engineering staff from Purdue University, and one of "laymen" who were "randomly selected as typical road users." The study was conducted in the following manner:

> Each rater was requested to drive over the pavement sections in a vehicle similar to the one that he normally drove. He could ride over the section at any speed desired but rating was not to be done during rain or inclement weather conditions. It was also stressed that the rater was to travel alone and work independently … [to avoid being] influenced by the opinions of others. (p. 25).

The results from main effects and interactions, based on an analysis of variance procedure, are shown in Table 3. Along with these results, a single linear regression was used to model the relationship between PSI and roughness (BPR roughometer measurements) for each of the three pavement types.

Table 3. Summary of Results from Nakamura and Michael (1963)

Significant	Not Significant
• Pavement Section (within Pavement Type) • Rater (within Rating Panel Type) • Interaction between Pavement Type and Rater (within Rating Panel Type)	• Pavement Type • Rating Panel Type • Interaction between Pavement Type and rating Panel Type • Interaction between Section (within Pavement Type) and Rating Panel Type • Interaction between Section (within Pavement Type) and Rater (within Rating Panel Type)

Janoff and Nick (1983)

In 1983, Janoff and Nick conducted noteworthy research for the Pennsylvania Department of Transportation (PennDOT). While copies of the original research report are difficult to obtain today, a few papers published from this research project are accessible. One paper, titled "Effects of Vehicle and Driver Characteristics on the Subjective Evaluation of Road Roughness," was published as part of an American Society of Testing and Materials (ASTM) conference proceedings. It summarized the results of an experiment that evaluated the effects of vehicle size, vehicle speed, and characteristics of the rating participants on the subjective evaluation of road roughness.

Passenger evaluators were used to test each of four hypotheses separately. In one experiment, two sets of 21 participants rated 11 test segments in Pennsylvania – one set was assigned to a 1982 Horizon and the second set was assigned to a 1982 Reliant.

In another experiment, two sets of 21 participants rated 34 test segments in Pennsylvania. Raters, who were assigned seats in one of two vehicles (a Horizon and a Reliant), were driven at two constant speeds; one set was driven at 40 kph (25 mph) and the second set was driven at 70 kph (45 mph). An analysis of variance (ANOVA) test concluded that the speed of the test vehicles did not have an impact on the average roughness measurement.[6]

This research also concluded that "novice" and "expert" panelists could be used to provide similar results and that the hypothesis that participant location ("regionality") did

[6] Speeds were held constant in some of the previous research because of the nature of the road roughness measurement systems. As Nick and Janoff explained, "The PennDOT standard operating procedure is to measure all roads at 40 mph (or 25 mph if absolutely necessary) because the PennDOT MRMs [Mays Ride Meters] have been calibrated at those speeds" (p. 7, 1983).

not have an impact on the average roughness measurement could not be rejected.[7] A summary of all four experiments is shown in Table 4.

Table 4. Summary of Hypotheses Tested by Janoff & Nick (1983)

Variable	Panel Number	Sites	Vehicle Type	Speeds	Null Hypothesis	Result
Panel Regionality	21 PA (1-a)	FL	compact	1 per site	No difference between the mean ratings for regionally different panels	Rejected
	21 FL (1-b)	FL	Compact	1 per site		
Vehicle Size	21 PA (1-a)	PA	Compact	1 per site	No difference between the mean ratings obtained for panels using different vehicles	Could not be rejected
	21 PA (2)	PA	Subcompact	1 per site		
Vehicle Speed	21 PA (1-a)	PA	Compact	1 per site	No difference between the mean ratings obtained for panels in vehicles driven at different speeds	Could not be rejected
	21 PA (3)	PA	Compact	6-8 site speeds changed		
Rater Training	21 PA (1-a)	PA	Compact	1 per site	No difference between the mean ratings made by trained and laymen panels	Could not be rejected
	21 PA (4)	PA	Compact	1 per site		

Source: Janoff & Nick (1983).

Nair, Hudson, & Lee (1985)

In 1985, research conducted at the University of Texas at Austin extended the work of Carey and Irick by validating the pavement serviceability concept. The primary purpose of this study was to use linear regression techniques to develop a new set of "serviceability prediction equations." These would model passenger ratings (i.e., PSI) as

[7] For the Janoff and Nick study, it is the authors' opinion that the conclusions based on the results of their statistical analysis could be interpreted in a slightly different manner. Whereas they concluded that "subjective ratings…will probably be different in a different area of the country." We suggest that the conclusion should have been that "subjective ratings cannot be considered to be the same in different areas of the country." The distinction is subtle but noteworthy when we compare results from one study to the next.

a function of roughness measures to reflect the "changed perceptions and expectation of the traveling public" (p. 170) and would be based on modern vehicles as well as newer, more accurate roughness measurement devices.[8] In this case, vertical acceleration was used as the roughness measure, and equations were developed for both flexible and rigid pavements.

Perhaps the most pertinent findings were the testing of significant variables. This study also used analysis of variance (ANOVA) techniques to test the significance of variables shown in Table 5. The results are shown in Table 6.[9]

Table 5. Variables Considered by Nair et al. (1985)

Pavement-Specific	Rater-Specific	Vehicle-Specific
• Roughness Measurement - PSR > 2.5 vs. PSR < 2.5 • Pavement Type - asphalt vs. PCC • Surface Texture - "coarse" vs. "fine" • Location - "rural" vs. "urban" • Maintenance - "unpatched" vs. "patched" • Surroundings - "scenic" vs. "unattractive" • Road Width - "narrow" vs. "wide" • Functional Class (Volume) - "low" vs. "high" • Lane Position - "inside" vs. "outside"	• Gender - male vs. female • Age - < 35 vs. > 35 years • Position in vehicle - front vs. rear • Profession - road profession vs. "layman" • Rater fatigue - "fresh" vs. "tired" • Time of day - day vs. night	• Type - car vs. van • Wheelbase - "short" vs. "long" • Speed - 30 mph vs. 50 mph

[8] The Nair et al. study used a Mays Ride Meter (MRM), a Walker Accelerometer Device (SIometer), and the 690D Surface Dynamics profilometer to collect roughness data and compare their correlations with panel ratings.

[9] Along with ANOVA results, simple comparisons plots of mean rankings of different variable-specific groups were made (e.g., mean male ratings vs. mean female ratings, or mean car ratings vs. mean van ratings) to explore differences among variables.

Table 6. Summary of Results from Nair et al. (1985)

Significant	Not Significant
• Surface roughness • Vehicle type • Vehicle size • Vehicle wheelbase length • Rater fatigue • Pavement type • Pavement maintenance	• Rater position in vehicle • Rater gender • Rater age • Time of day • Rater profession, technical knowledge • Rater function in vehicle • Vehicle speed • Pavement surface texture • Pavement location • Road width • Surroundings

This study had other interesting findings in terms of user acceptability. The only problem is that all results were in terms of PSRs, as shown in Figure 4.

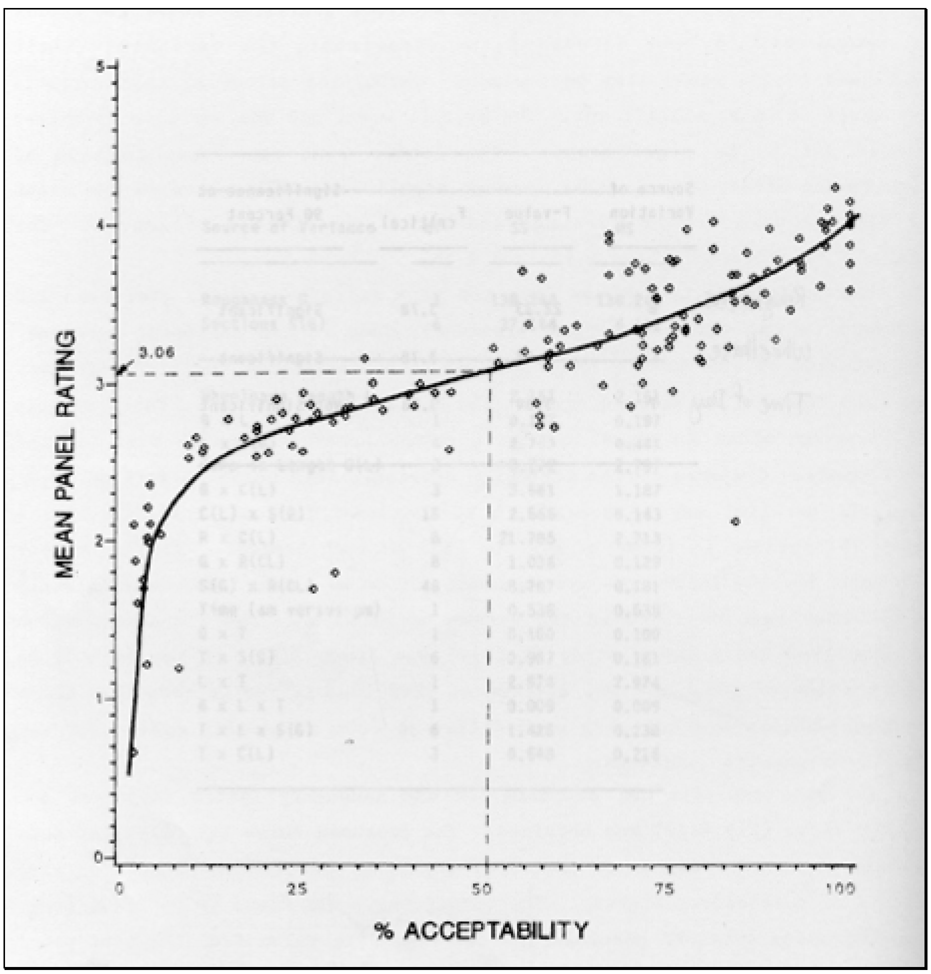


Figure 4. PSR Acceptability
Source: Nair et al., 1985, p. 108.

Other results regarding acceptability include the following:

- "Sections with high PSRs (greater than 3.5) were rated to be acceptable unanimously by the panel" (p. 81
- "Up to 88 percent of the variation in PSR can be explained by the roughness variables" (p. 81).
- "This study yet further attests to the serviceability-performance concept in general and to the validity of using road profile measurements to predict PSRs and to obtain indices of serviceability in particular" (p. 163).

Garg et al. (1988)

Garg et al. produced somewhat similar results by creating a separate set of PSI prediction equations using linear regression. Like Nair et al., the researchers also did some significance testing, but they did not document any of these findings:

> A number of subjects' personal and background variables had statistically significant effects on both subjective measure of road roughness and the acceptability measures. However, these variables were not found to be of practical significance as they explained very little variation in either road roughness or acceptability measures. (p. 282, Garg et al., 1988).

Fernando and Lee (1999)

A current study conducted by the Texas Transportation Institute (TTI) has not been completed, but preliminary results compare to findings from the work done in 1985. The current study used 28 evaluators, consisting of TTI employees, Texas DOT employees, and Texas A&M students, to evaluate ride quality over 63 pavement segments. These subjects were grouped into pairs and used as passengers in test vehicles. The subjects were driven at a constant speed of 50 mph over pre-determined pavement segments with measured roughness indices while being asked to evaluate overall ride quality using the AASHO/Weaver scale. (See Appendix E for details on different rating scales.) Researchers assigned a numerical value "in tenths of a rating point" based on the mark that the rater made on the rating scale.

Over 3,000 evaluations were collected. Using analysis of variance (ANOVA), this study uncovered some basic relationships between PSR and IRI. Factors that were tested and deemed to be significantly correlated with passenger roughness ratings included section roughness, vehicle type, the individual rater, and pavement type. The study concluded that raters had a tendency to rate PCC sections to be rougher than asphalt sections. It also acknowledged that it was not ultimately trying to identify the relationship between ride quality and pavement or vehicle factors, "While these findings suggest that relationships for evaluating ride quality should include pavement and vehicle factors, the evaluation of such relationships would require a much larger effort than planned for the current project" (p. 33).

Kuemmel et al. (2001a,b,c)

A series of reports, titled "Public Perceptions of the Midwest's Pavements," documented a five-year, three-phased effort supported by pooled funds from Wisconsin, Iowa, and Minnesota. The goal of the project was to "assess the public's perceptions of pavement improvement strategies, develop customer-based thresholds of satisfaction as related to the Departments' physical indices, such as pavement ride and condition." (The project was conducted as three independent studies in each of the three states, and a separate report was written for each phase.)

The first phase of the study (Phase I) involved the analysis of 381 responses within a series of six focus groups in each state to better evaluate each state DOT's pavement policies. Phase II involved the administration and analysis of 400 statewide telephone interviews in each state to gather information about improvement policy trade-offs. (The surveys included questions on satisfaction, trust, improvement issues, construction delays and detours, and trade-offs regarding improvement strategies.)

In Phase III, participants were recruited by phone and asked to drive at their leisure (presumably using their own vehicle) over selected rural highway segments ("within 10 minutes drive time of a city of 500 population or more"). Participants were called back at a later time and asked general questions about their satisfaction with the segment. Participants received $10 for their time and any related expenses – if they completed the follow-up phone survey within approximately one week. In six months, over 450 highway segments were selected, and 2,300 surveys were completed in the three states.

While the details on the data collection were limited, it appears that the data were collected under unrestricted conditions, insofar as no effort was made to control the conditions during which drivers evaluated the segment. No aspect of this phase of the data collection appears to have been controlled other than the selection of the test section.

Using these data, researchers tried to determine the roughness and distress levels that are tolerated by the public. This issue was investigated by relating IRI, along with pavement condition indices (PCI), to the cumulative percentage of respondents who agreed with each of three "threshold" questions related to satisfaction:

1. "I am satisfied with the pavement on this section of highway" ("satisfied")
2. "The pavement on this stretch of highway is better than most of the stretches of state highways I've driven in Minnesota" ("better than most")

3. “The pavement on this stretch of highway should be improved” (“improve”).

From these questions, researchers attempted to determine the IRI value at which 70 percent of drivers would be satisfied with a given section of highway, shown in Table 7. This study is one of the few studies that identified IRI acceptability levels.

Table 7. IRI Acceptability Levels in Kuemmel et al. (2001)

Sample Size	Pavement Type			
	PCCP (N = 240)	**ACP (N = 171)**	**Composite (N = 203)**	**Total (N = 614)**
10%	203 in/mi (3.2 m/km)	158 in/mi (2.5 m/km)	171 in/mi (2.7 m/km)	184 in/mi (2.9 m/km)
20%	158 in/mi (2.8 m/km)	120 in/mi (1.9 m/km)	120 in/mi (1.9 m/km)	158 in/mi (2.5 m/km)
30%	158 in/mi (2.5 m/km)	108 in/mi (1.7 m/km)	89 in/mi (1.4 m/km)	127 in/mi (2.0 m/km)
40%	146 in/mi (2.3 m/km)	76 in/mi (1.2 m/km)	76 in/mi (1.2 m/km)	108 in/mi (1.7 m/km)
50%	120 in/mi (1.9 m/km)	63 in/mi (1.0 m/km)	70 in/mi (1.1 m/km)	76 in/mi (1.2 m/km)
60%	114 in/mi (1.8 m/km)	51 in/mi (0.8 m/km)	57 in/mi (0.9 m/km)	63 in/mi (1.0 m/km)
70%	70 in/mi (1.1 m/km)	44 in/mi (0.7 m/km)	-	-

Source: Kuemmel et al., 2001a, p. 12.

Table 8 summarizes the information from all of the literature sources described in this chapter.

Table 8. Summary of Previous Research

Authors (Year)	Location	Number of Participants	Number of Test Segments	Test Vehicles	Roughness Measurement System	Analysis Methodology
Carey & Irick (1960)	Ill., Ind., Minn.	100	(74 flex, 49 rigid)	N/A	BPR Roughometer	Linear regression
Nakamura & Michael (1963)	Indiana	30	60 (19 flex, 22 comp, 19 rigid	uncontrolled	BPR Roughometer	• ANOVA • Linear regression
Scrivner & Hudson (1964)	Texas	12	43 (all flex)	(N/A)	CHLOE Profilometer	Linear regression
Karan et al. (1979)	Ontario (Canada)	8	55	midsize	Automatic Road Analyzer (ARAN)	• ANOVA • Linear regression
Janoff & Nick (1983)	Penn. & Florida	63	65	• subcompact • compact	• Mays Ride Meter • Profilometer	ANOVA
Janoff et al. (1985)	Ohio	36	81 (25 flex, 34 comp, 22 rigid)	N/A	• Mays Ride Meter • Profilometer	Linear regression
Nair et al. (1985)	Texas	20 (15 raters, 5 drivers)	179 (131 flex, 48 rigid)	• subcompact • midsize • van	• Mays Ride Meter • SIometer • 690D Surface Dynamic Profiler	• ANOVA • Linear regression

(Table 8. Summary of Previous Research continued)

Moore et al. (1987)	Kansas	24	103	midsize	Mays Ride Meter	Linear regression
Garg et al. (1988)	Wisconsin	50	32 (21 bit, 11 PCC)	midsize	• Wisconsin Roadmeter • Michigan Profilometer	• Linear regression
Arterbrun & Suprenant (1990)	Colorado	24	69	"cars"	• Mays Ride Meter	• ANOVA • Linear regression
Ward et al. (1993)	Indiana	68	66 (42 bituminous, 24 concrete)	midsize	• FHWA ProRut • Model F Cox Roadmeter • Indiana RIP Profiler	Linear regression
Fernando & Lee (1999)	Texas	28	63 (41 AC, 12 PCC)	• midsize car • minivan • van • pickup		ANOVA
Kuemmel et al. (2001)	Iowa Minnesota Wisconsin	N/A	N/A	uncontrolled	N/A	

Notes: In all cases, all participants were passenger raters – unless noted.
The notation (N/A) indicates that the information was not available.

EXTENDING THE CURRENT BODY OF KNOWLEDGE

This research took into account the work done by other researchers and extended the current body of knowledge. Like the previous research, this research continued to move beyond physical indices and supported the premise that physical indices alone provide an incomplete characterization of road roughness in the eyes of the driving public.

From a data perspective, this study combined roadway-specific, vehicle-specific, and individual-specific data. Unlike other research dating back to the 1960s, this research was able to take advantage of modern roughness measurement systems and other contemporary, pavement management system data. Additionally, it drew its conclusions entirely from a set of drivers, not passengers, as in most of the previous research, and allowed for as much normal driving behavior as possible by participants. It also drew data from four distinct vehicle types.

Above all, this study moved beyond the traditional performance-serviceability concept and the accompanying panel rating methods originally employed during the AASHO Road Test. Unlike traditional pavement-serviceability ratings, which are collected with such precision that they are regarded as a stand-alone "performance measure," the individual roughness rankings in this study were designed to focus more on capturing variability among test segment evaluations. These data were easily recorded by researchers in the passenger seat of the test vehicle as participants drove. While the discrete choice data collected in this experiment were not as precise (or as comparable) as the continuous scale assumed with the passenger ratings, these data still allowed for more rigorous statistical analysis than in the past. While most past analyses were limited to linear regression or ANOVA, this study took advantage of econometric probability models that identified significant variables as well as their relative impact on roughness ranking and acceptability – all of which could be valuable for making rehabilitation decisions.

CHAPTER 3. DATA COLLECTION

Data in this study originated from three primary sources: 1) a preliminary mailout/mailback survey, 2) an in-vehicle study, and 3) the Washington State Department of Transportation and its Pavement Management System (WSPMS). In this chapter, each of these data sources is discussed in detail. Together these data sources were combined to produce a data set that contained individual driver information, the in-vehicle driver evaluations of specific roadway segments, and physical characteristics of those specific roadway segments. However, the selection of this data set may not represent all Washington State locations.

PRELIMINARY MAILOUT/MAILBACK SURVEY

A simple mailout/mailback survey was sent to over 2,500 registered vehicle owners in Seattle and the surrounding area. The purposes of the preliminary mailout/mailback survey were two-fold. The first was to obtain an initial pool of potential participants who normally drive on Seattle-area highways; the second was to assess who would be willing to participate in an in-vehicle study. To accomplish this goal, a random sample of vehicle license plates was collected from vehicles entering and exiting SR 520 near the University of Washington. Over 2,800 license plates were gathered during various hours of the day and random days of the week in the fall of 2000.

Addresses corresponding to the collected license plates were obtained through the Department of Licensing, and the survey was sent to the registered owner of the vehicle. The simple, two-page survey contained 26 questions and included a postage-paid business reply envelope. It was designed to get the public's general opinion of local area highways and to determine whether the respondent would be willing to participate in an in-vehicle study at a later time. Respondents were asked to provide contact information (name, address, and phone number) so they could be contacted at a later time if they were interested in participating in the in-vehicle study. The preliminary survey instrument is included in the appendix, and the response rates from the mailout/mailback survey are presented in Table 9.

A fraction of the collected license plates were invalid and did not yield a mailing address. As a result, a total of 2,570 surveys were mailed, of which 621 surveys were returned. Of the 621 surveys received, 529 were valid responses and 92 were returned as "return to sender." After invalid mailing addresses were taken out, the response rate of the preliminary survey was over 21 percent, as shown in Table 9, which is comparable to other mailout/mailback surveys with response rates of 33 percent (Khattack et al., 1993) and 15 percent (Ng et al., 1995), among others.

Table 9. Preliminary Survey Response Rates

Result	Number	Percent
Returned *With* Contact Info	173	7.0%
Returned *Without* Contact Info	356	14.4%
TOTAL	529	21.4%

Besides identifying a pool of potential participants for the in-vehicle phase of the study, the preliminary survey collected general opinions about road roughness around the Seattle area. More importantly, the survey also collected socioeconomic data, including vehicle use and ownership information.

IN-VEHICLE STUDY

The in-vehicle data collection was relatively simple. Drivers who wrote down their contact information on the preliminary survey were contacted in the late spring of

2001 with a simple one-page letter and instructed to telephone the University of Washington if they were still interested in participating. (This letter was also necessary, in part, because nearly six months had passed between the time that participants had initially provided their contact information on the preliminary survey and the time at which the in-vehicle data collection took place.) The one-page letter, included in the appendix, provided a little more information about the in-vehicle experiment and included the promise of $50 for less than two hours of their time.

When contacted on the phone, participants were told more about the in-vehicle data collection study, and any questions they had about the study were answered. Once they agreed to participate, a time was scheduled for the in-vehicle driving experiment, and they were given directions to meet a researcher at one of two randomized starting locations. The methodology for the study is explained in the next section. The study was conducted in the early summer of 2001.

Survey Instrument

At the randomized starting location, participants were simply told that the Department of Transportation wanted their opinions on road roughness for highways in the area. They were shown a general map (see Figure 5) of the highway segments and told that they would be driving over 40 predetermined highway test "segments." As they drove over the test segments, they would be asked two questions:

1. "How would you rank the roughness of the road on a scale from one to five – with one being the smoothest or the best and five being the roughest or the worst?"
2. "Is this level of roughness acceptable to you?"

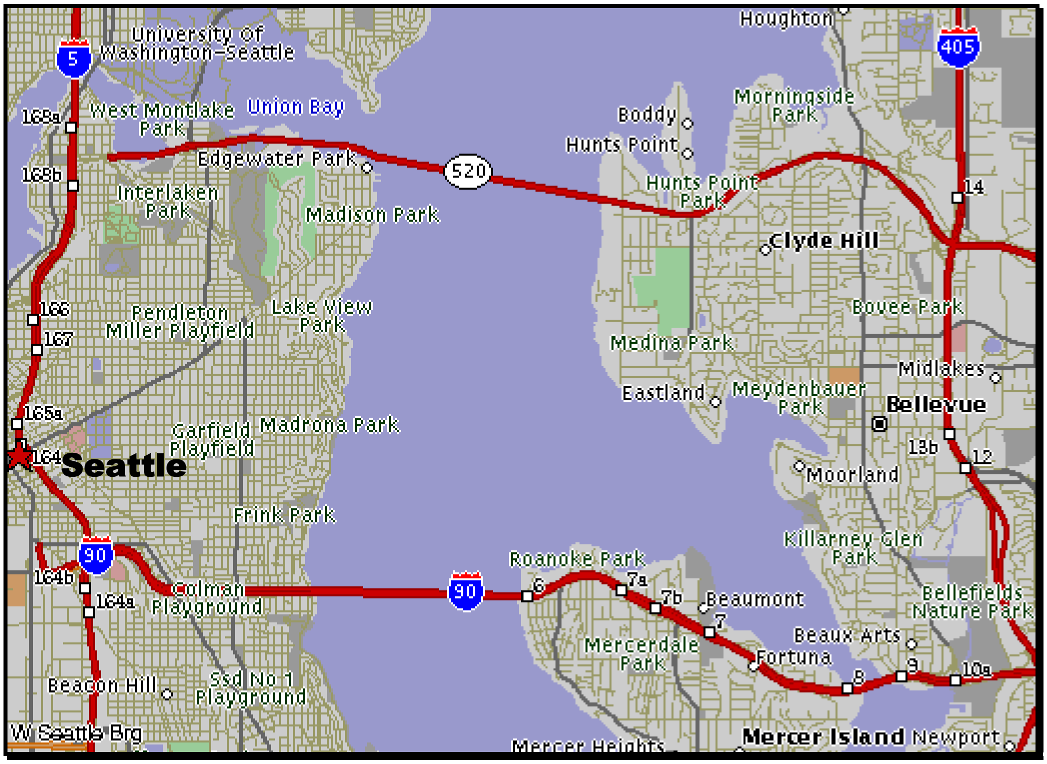


Figure 5. General Map of Study Area

Participants were not provided any more instructions. They were told that during the driving experiment they would be notified when each segment started and ended and could provide their response to the two questions discussed earlier at any point during the segment. They were told that their roughness rankings needed to be in the form of whole numbers (i.e., "no use of fractions or decimals") and that their response to the second question was a simple "yes" or "no" answer. They were not given any other explicit instructions – except where to drive. If they asked how "roughness" or "acceptable" were defined, they were told to use their own judgment and to rank the roughness of the road in comparison to other roads in the state. In other words, participants were not given preset evaluation criteria and were forced to develop their own roughness ranking criteria. They were also instructed to drive in a way that was "consistent with their regular driving behavior." They were instructed to drive at a speed that they felt most comfortable driving. The purpose of this stipulation was to capture as much of the participant's normal driving behavior as possible with the hopes that the "research setting" would not affect their responses.

Study Area and Route Selection

Route selection was an important but challenging step in the data collection process because it literally represented the foundation on which all the results were to be based. The routes selected needed to meet certain criteria. For convenience and for comparison to the JLARC study, the roadway segments needed to be on state-managed highways and in the Seattle area, which participants with varying driving abilities could easily and willingly access. In addition, the routes within the study area needed to be easily and safely traversable by drivers of varying abilities and in a reasonable amount of time (less than two hours). The pavement segments over which the participants drove needed to have consistent and homogeneous attributes, such as consistent design geometrics, pavement type, and pavement condition. At the same time, it would be ideal if the test segments sampled had various ranges of known physical indices. Therefore, it was important for the segments to be located where physical indices had been documented in the Washington State Pavement Management System (WSPMS). Finally, a useful feature of each pavement section was that distinguishable and easily identifiable roadway attributes separate it.

On the basis of all of these requirements, a 25-mile circular loop on I-5, I-90, I-405 and SR 520 around the Seattle-Bellevue area was selected. This loop was chosen for many reasons. First, it was easily accessible by many participants because it was close to the Seattle and Bellevue areas and sampled four major facilities in the area – all of which have unique attributes. To expand the data collection, the loop was driven twice – once in a clockwise direction and once in a counterclockwise direction. There were 20 test segments on each loop (40 total segments for both loops). Table 10 summarizes the IRI measurements on the four sampled facilities by increasing and decreasing state milepost.

Table 11 organizes the IRI measurements by loop direction. A critical analysis of these IRI measurements is discussed in Chapter 5.

Table 10. Summary of IRI Measurements on Sampled Facilities by Milepost

Facility	Increasing Milepost (MP)				Decreasing Milepost (MP)			
	Min	Max	Mean	Standard Deviation	Min	Max	Mean	Standard Deviation
I-5	164 in/mi (2.59 m/km)	220 in/mi (3.47 m/km)	185 in/mi (2.92 m/km)	15 in/mi (0.230 m/km)	106 in/mi (1.68 m/km)	150 in/mi (2.36 m/km)	128 in/mi (2.02 m/km)	15 in/mi (0.231 m/km)
I-90	82 in/mi (1.30 m/km)	182 in/mi (2.88 m/km)	112 in/mi (1.77 m/km)	20 in/mi (0.321 m/km)	73 in/mi 1.15 m/km	120 in/mi (1.89 m/km)	97 in/mi (1.53 m/km)	4 in/mi (0.058 m/km)
I-405	163 in/mi (2.58 m/km)	271 in/mi (4.28 m/km)	224 in/mi (3.53 m/km)	32 in/mi (0.506 m/km)	62 in/mi 0.98 m/km	127 in/mi (2.00 m/km)	90 in/mi (1.42 m/km)	16 in/mi (0.258 m/km)
SR520	77 in/mi (1.22 m/km)	191 in/mi (3.01 m/km)	112 in/mi (1.77 m/km)	34 in/mi (0.532 m/km)	56 in/mi (0.89 m/km)	208 in/mi (3.28 m/km)	102 in/mi (1.61 m/km)	60 in/mi (0.949 m/km)

Notes: For I-5 and I-405, increasing mileposts represent the northbound direction, while decreasing mileposts represent the southbound direction. For I-90 and SR 520, increasing mileposts represent the eastbound direction, while decreasing mileposts represents the westbound direction.

Table 11. Summary of IRI Measurements on Sampled Facilities by Direction

Direction	Min	Max	Mean	Standard Deviation
Clockwise	62 in/mi (0.98 m/km)	220 in/mi (3.47 m/km)	113 in/mi (1.78 m/km)	113 in/mi (1.78 m/km)
Counterclockwise	56 in/mi (0.89 m/km)	271 in/mi (4.28 m/km)	132 in/mi (2.08 m/km)	70 in/mi (1.10 m/km)

Notes: The clockwise direction consisted of eastbound SR520, southbound I-405, west-bound I-90, and northbound I-5, but not necessarily in that order. The counterclockwise direction consisted of southbound I-5, eastbound I-90, northbound I-405, and westbound SR 520, but not necessarily in that order.

The decision to repeat the loop in the opposite direction was deliberate. It was done for both logistical and experimental reasons. It allowed drivers (and the researcher) the ability to return to the starting location halfway through the experiment if any unforeseen problems prevented the experiment from being completed in less than two hours (e.g., a major traffic accident). A loop was also selected because it allowed the participant to drive on each facility twice (albeit in opposite directions) to increase driver comfort with the general route.

For logistical reasons, most test segments were located in the "slow," rightmost lane. The "slow" lane is easily traversable by individuals of varying driving abilities, regardless of the amount of congestion present. The number of lane changes required during the experiment was minimized for increased safety and simplicity. Drivers were told that most of the time they would be driving in the rightmost lane "unless otherwise instructed." A few test segments were not located in the "rightmost" lane and were selected as a result of the way some of the facility interchanges were designed. After changing between facilities, the adjoining "on-ramp lane" was not always in the rightmost lane, and in this situation it was easier for some test segments to be located in the adjoining lane to prevent unnecessary lane changes. Regardless of their driving ability, as will be discussed in the next section, all drivers traversed the same test sections in the same lanes.

In-Vehicle Data Collection

As noted earlier, participants were surveyed about their perceptions of road roughness as they drove over the same 40 local highway test segments, and they were instructed to rank the roughness of the road and to decide whether the roughness was acceptable as they drove over each pavement segment. While participants drove over each test segment and answered the two evaluation questions, a researcher located in the passenger seat collected other data that the driver was not informed about until after the experiment was complete. These data included the following:

- Noise: A decibel meter was in the cabin to monitor the noise levels (dB) inside the interior of the vehicle.[10]
- Vehicle Speed: Vehicle speed was recorded from the in-vehicle speedometer at the time driver responses were recorded.
- Weather and Pavement Moisture: Although data were collected to note any inclement weather and pavement moisture, virtually all data collection occurred in the summer during dry weather conditions.

The starting location, weather conditions, and vehicle type were noted before the driving experiment. A discussion of starting location and vehicle type follow.

Randomization of Starting Locations

Before the participant was scheduled a meeting time, the starting location was randomized to minimize the effects of driver fatigue. One start location was University Village, near the junction of I-5 and SR 520, and the second start location was the Eastgate Park and Ride (P&R), near the junction of I-90 and I-405 in Bellevue, as summarized in Table 12 and shown in Figure 6. When starting at the Eastgate P&R and driving in the *clockwise* direction, the participant experienced (in order) I-90, I-5, SR 520, and I-405; if instructed to drive in the *counterclockwise* direction from the Eastgate

[10] Noise levels were recorded using the C-weighted decibel scale, dBC. C-weighted levels are not used as frequently as A-weighted levels in transportation research, but they may be preferable in evaluating sounds whose low-frequency components are responsible for secondary effects such as the shaking of a vehicle, window rattle, perceptible vibrations, or other factors that can cause annoyance while driving.

P&R, the participant experienced (in order) I-405, SR 520, I-5, and I-90. As will be shown in the results, the differences observed in drivers perceptions of roughness were not associated with start location.

Table 12. Summary of Start Locations

Place	Location	Initial Driving Direction
Start Point 1	University Village	Clockwise
Start Point 2	University Village	Counterclockwise
Start Point 3	Eastgate P&R	Clockwise
Start Point 4	Eastgate P&R	Counterclockwise

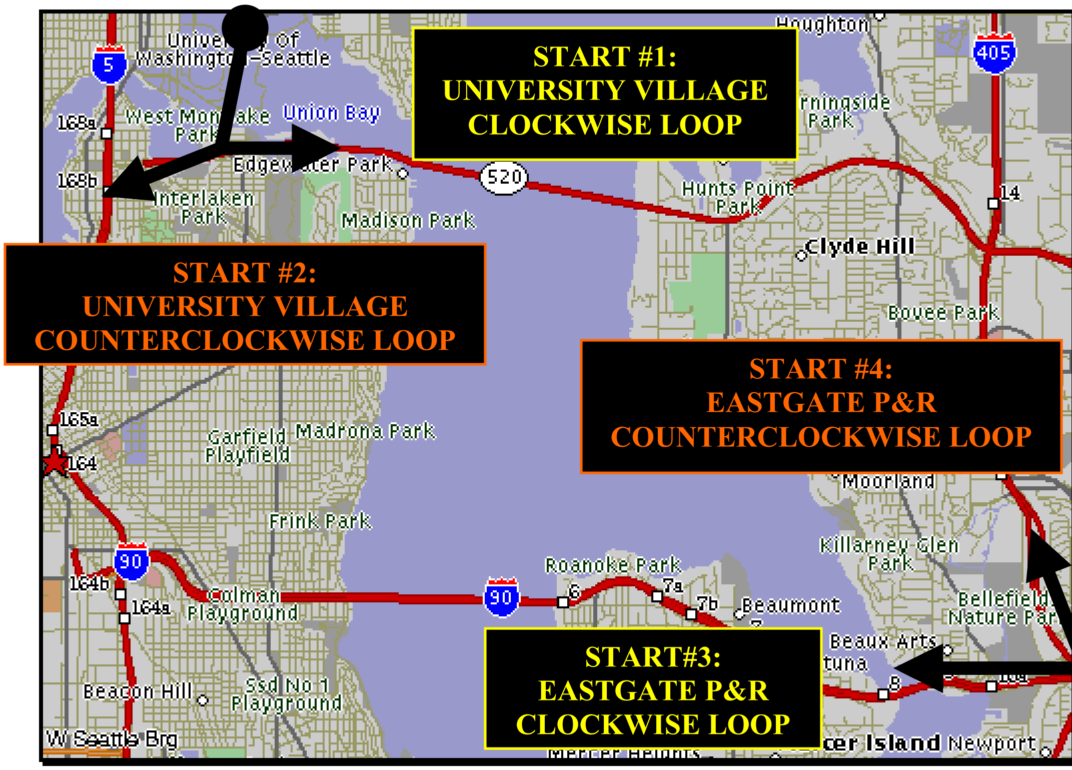


Figure 6. Map of Study Area Indicating Randomized Starting Locations

Selection of Test Vehicles

Given that vehicle type could have an impact on the perception of road roughness, drivers were also assigned to different vehicle types. Vehicle type was thought to be important not only because of the characteristics of the vehicle but also because of the way the vehicle interacted with the roadway. Four types of vehicles were used in the study. The number of participants for each start location and vehicle type is summarized in Table 13.

Table 13. Summary of Driving Participants by Start Location and Vehicle Type

Vehicle Type	Vehicle	Start Location	Participants
Sedan	2001 Kia Optima	Start Location 1	1 (5.6%)
		Start Location 2	4 (22.2%)
		Start Location 3	7 (38.9%)
		Start Location 4	6 (33.3%)
SUV	2001 Jeep Grand Cherokee	Start Location 1	5 (31.3%)
		Start Location 2	5 (31.3%)
		Start Location 3	2 (12.5%)
		Start Location 4	4 (25.0%)
Pickup	2001 Ford Ranger	Start Location 1	3 (25.0%)
		Start Location 2	4 (33.3%)
		Start Location 3	3 (25.0%)
		Start Location 4	2 (16.7%)
Minivan	2001 Ford Windstar	Start Location 1	3 (30.0%)
		Start Location 2	3 (30.0%)
		Start Location 3	2 (20.0%)
		Start Location 4	2 (20.0%)

There was an unbalanced number of participants in each cell for a couple of reasons. Participants were randomly assigned a start location and a vehicle type independently, and even though each participant had an equal probability of being assigned to a specific vehicle type at a specific start location, the final distribution was not uniform. Second, the test vehicles were rented and used one at a time for cost considerations. As the in-vehicle study data collection progressed, scheduling participants became more and more difficult as the pool of willing participants grew smaller, and there were a greater number of missed ("no show") appointments. All of these reasons, coupled with time and cost constraints, resulted in uneven sampling at different start locations and with different test vehicles.

PHYSICAL SEGMENT-SPECIFIC WSDOT DATA

The last data source was the WSDOT and its pavement management system.[11] The Washington State Pavement Management System provided physical data on each segment and was instrumental in the segment selection process by providing valuable geometric information that was helpful in selecting test segments. In particular, the database contained information about terrain, shoulder width, number of lanes, and roadway width, all of which were required to be homogeneous within each segment as part of the selection criteria.

Along with the WSPMS, the WSDOT State Route Viewer (SRView) was used in the pavement selection process. This database of images allowed us to match milepost demarcation with notable landmarks and features on the roadway. A booklet of images was created with SRView that showed pictures of the location where each segment started and ended.

Along with the segment selection process, the WSPMS also contributed to the final data file used in the analysis. Each test segment was characterized by type of surface (PCC or AC), a list of prior rehabilitations of each segment, and whether it was a bridge segment. From this information, the age of the surface layer was determined. The WSPMS also contains data on physical indices that WSDOT uses to evaluate pavements and prioritize rehabilitation. It contains measurements for the pavement ride (IRI), rutting, and the pavement structural condition (PSC), all of which are useful pavement

[11] See Sivaneswaran et al. (2000) for details.

evaluation criteria.[12] Even though the WSPMS contains IRI data, a WSDOT data collection crew measured the roughness profile using the pavement distress identification vehicle (shown in Figure 7) and obtained updated IRI values for the test segments in lieu of the IRI values in the WSPMS to ensure accuracy of the data at the time of the study.

Figure 7. WSDOT Pavement Distress Identification Van

Source: WSDOT, 2001.

SUMMARY OF DATA SOURCES

In total, 56 driving participants each evaluated 40 highway segments and produced nearly 2,240 unique "observations." (Actually, there were 2,180 valid "observations" because of missing or incomplete data for some participants.) For each "observation," there were associated quantitative socio-demographic data for the individual from the preliminary survey, physical segment-specific data from the WSPMS, and in-vehicle survey data from the driving experiment. Tables 14, 15, and 16 summarize the relevant data collected from each data source.

[12] The pavement structural condition (PSC) has values ranging from 100 (excellent pavement condition) to zero (completely deteriorated pavement).

Table 14. Summary of Relevant Data from Preliminary Survey

Data Description	Type
Regular Seattle-area commuter	Binary
Respondent gender	Categorical
Respondent's marital status	Categorical
Respondent age	Categorical
Respondent annual household income	Categorical
Respondent highest level of education	Categorical
Model year of vehicle most often driven	Discrete
Frequency of seatbelt usage	Categorical
Number of accidents in the past 5 years	Discrete
Type of vehicle most often driven	Categorical
Number of vehicles in household	Discrete
Number of people in household	Discrete
Number of children under age 6 in household	Discrete
Number of children 6 to 16 in household	Discrete
Number of people in household working outside the home	Discrete
Average vehicle occupancy	Categorical
Frequency of use of I-5	Ordinal
Frequency of use of I-90	Ordinal
Frequency of use of I-405	Ordinal
Frequency of use of SR520	Ordinal

Table 15. Summary of Relevant Respondent Data from In-Vehicle Survey

Data Description	Type
Ranking of segment roughness on 1 (smooth) to 5 (rough) scale	Ordinal
Is the roughness acceptable? (yes/no)	Binary
Test vehicle type (sedan, pickup, sport utility, minivan)	Categorical
Noise measurements inside vehicle (dB)	Continuous
Test vehicle type	Categorical
Speed measurements of vehicle during in-vehicle response	Continuous
Estimated level of service (LOS) during in-vehicle response	Ordinal

Table 16. Summary of Relevant Pavement Data from Test Segments

Data Description	Type
Start milepost	Continuous
End milepost	Continuous
IRI (measured by WSDOT)	Continuous
Segment surface type	Categorical
Segment has exposed aggregate present (yes/no)	Binary
Segment has expansion joints present (yes/no)	Binary
Segment has rutting present (yes/no)	Binary
Segment has patches (slab replacement) present (yes/no)	Binary
Segment is located on bridge or elevated structure (yes/no)	Binary
Left shoulder width (feet)	Continuous
Right shoulder width (feet)	Continuous
Road width (feet)	Continuous
Number of lanes	Discrete
Rutting	Continuous
Pavement structural condition (PSC)	Continuous

CHAPTER 4. DATA ANALYSIS

A couple of statistical analysis techniques can be used to link various vehicle type, roadway characteristics, and driver characteristics with perceived road roughness. This chapter presents the fundamental theory and justification for selecting the statistical techniques to analyze these data.

The purpose of the analysis was two-fold: to identify differences in perceived roughness and to identify the factors that are most significantly associated with the differences in driver perception. These goals were accomplished by extending beyond simple tests of statistical significance and by taking advantage of more rigorous statistical techniques. In particular, the combined data set lent itself well to both analysis of variance techniques to identify significant differences in start location and vehicle type, and an ordered econometric probit (or logit) model to identify the factors that were most significantly associated with the differences in drivers' roughness perception.

ANALYSIS OF VARIANCE (ANOVA)

Analysis of variance (ANOVA) is used to test the hypothesis that several means are equal. This analysis technique is constructive because it provides a method to determine whether the observed differences among groups of means are equal, or the underlying factors have a potentially statistically significant impact on the response. This technique assumes that the continuous responses are independently and normally distributed, and given the nature of these data, certain precautions must be taken to ensure that these assumptions are not violated. For the case in which two factors are taken into account but there is no interaction among them, the general two-factor ANOVA model can be written as follows:

$$Y_{ij} = \mu_{..} + \alpha_i + \beta_j + \varepsilon_{ij} \qquad \text{Equation 1}$$

where Y_{ij} is the response, $\mu_{..}$ is the overall mean for the response over populations, α_i and β_j represent the effects of two independent treatments, and ε_{ij} is the error term. (See Neter et al., 1990 or Hicks, 1993 for a more detailed explanation.)

In this case, ANOVAs were conducted to test the null-hypothesis that the mean vehicle speed during the in-vehicle survey was equal, while controlling for other (independent) factors, such as vehicle type, pavement type, or starting point.

ORDERED PROBABILITY MODELS

Once some significant explanatory variables had been identified, their relationship with perceived road roughness could be explored. These explanatory variables could provide some quantifiable validation of hypotheses that linked differences between perceived road roughness and actual road roughness (i.e., measured IRI values). The relationship between explanatory variables and perceived road roughness could also provide some hypothetical justification to support the theory that additional measurements (other than roughness indices) might be warranted to fully explain perceived road roughness by drivers.

Most often, the relationship between a variable of interest and set of explanatory variables is established through linear regression. The problem with using linear regression, however, is that the dependent variable is assumed to be continuous. Perceived driver roughness is not a continuous variable; it is a ranked, ordinal response.

Linear regression would not be appropriate because the dependent variable in this case reflected a ranking, so that the difference between a 3 and a 4 could not be treated as equivalent to the difference between 2 and 3, for example (Greene, 1993).

To accurately model the discrete and ordinal properties of the roughness rankings, statisticians would recommend that "probability models," such as the logit or probit, be used instead of linear models (e.g., regression) because they allow a mixture of categorical and continuous independent variables with respect to a categorical (nominal or ordinal) dependent variable. Probability models try to predict the likelihood of an event occurring rather than modeling the actual response and are better suited to treat the case of dichotomous or polytomous dependent variables.

Logit and Probit Models

Two of the most common probability models are the logit and probit models. These two models are very similar in nature and design and can be written in a general form as follows (Greene, 1993):

$$y^* = \sum_{k=1}^{K} \beta_k x_k + \varepsilon \qquad \text{Equation 2}$$

where:

y^* is a stochastic, unobserved response

β_k are estimable model parameters

x_k are measurable factors (i.e., driver data, roadway characteristics, etc),

ε are unobservable disturbances that are assumed to be normally distributed across observations.

Instead of y^*, we actually observe y among the J number of ordered categories, such that

$$y = 0 \text{ if } y^* \leq 0 \qquad \text{Equation 3}$$

$y = 1$ if $0 < y^* \leq \mu_1$

$y = 2$ if $\mu_1 < y^* \leq \mu_2$
.
.
.
$y = J$ if $\mu_{J-1} < y^*$.

The unknown threshold parameters (μ_j) are estimated on the basis of the model parameters (β_j), such that probabilities are obtained. In the case of the probit model, the probabilities are calculated by assuming a cumulative normal distribution function, denoted by $\Phi(\cdot)$, so that:

$$\text{Prob}(y = 0) = \Phi(-\boldsymbol{\beta}'\mathbf{x}), \qquad \text{Equation 4}$$

$$\text{Prob}(y = 1) = \Phi(\mu_1 - \boldsymbol{\beta}'\mathbf{x}) - \Phi(-\boldsymbol{\beta}'\mathbf{x})$$

$$\text{Prob}(y = 2) = \Phi(\mu_2 - \boldsymbol{\beta}'\mathbf{x}) - \Phi(\mu_1 - \boldsymbol{\beta}'\mathbf{x})$$
.
.
.
$$\text{Prob}(y = J) = 1 - \Phi(\mu_{J-1} - \boldsymbol{\beta}'\mathbf{x}).$$

Because the left-hand sides of the equations are probabilities, they must lie between 0 and 1 for all values of x and for all parameter values. (The primary difference between the logit and probit models is that probabilities in the logit model are calculated by assuming a cumulative logistic distribution.)

In this case, when asked to rank the roughness of a given road segment, respondents of the in-vehicle survey had their own intensity of feelings regarding road roughness, which could vary depending on the measurable factors (x) (e.g., vehicle type, driver characteristics, speed, noise) and unobservable factors (ε). In principle, the respondents had an opinion about the road roughness, but they were only given five possible answers, so they had to choose the roughness ranking (y) that most closely represented their true opinion (y^*). For example, while a respondent may have wanted to rank the roughness of a specific segment as a 2.5, he/she was forced to rank it as a 2 or a 3, which was observed. This inherent decision-making process can sometimes be difficult with rank responses, especially when respondents are not permitted to use fractions or decimals in their response.

Each model presented in the next chapter contains descriptions of each statistically significant explanatory variable, the corresponding estimated model parameter β, and the corresponding t-statistic. This t-statistic indicates the significance of the explanatory variable and is computed from β/σ, where σ is the standard error of the explanatory variable and β is the estimated model parameter, as previously noted. Also included are the three estimated threshold values (μ), which bound the probability of observing the five roughness ranking categories. (See Greene (1993) or Kennedy (1992) for a more detailed explanation of probability models.)

Goodness-of-Fit Statistics

A goodness-of-fit measure for this model is the likelihood ratio, LR, which can be calculated as follows:

$$LR = -2[\ln L(C) - \ln L(0)] \qquad \text{Equation 5}$$

where lnL(C) is the log-likelihood at convergence, and lnL(0) is the initial log likelihood, which is calculated with only a constant term and no variables. The log-likelihood at convergence is the maximized value of the log-likelihood, taking all explanatory variables into consideration.

The LR statistic is chi-square (χ^2) distributed with degrees of freedom equal to the number of restrictions in the converged model. If the calculated likelihood ratio statistic is higher than the tabulated chi-square value for the given degrees of freedom, then it can be assumed that the variables make a statistically significant contribution to the initial log-likelihood model. (The null hypothesis that the slope coefficients of the tested variables are jointly equal to zero can be rejected.) In general, if two models have the same number of parameters, the model with the higher likelihood ratio statistic is better specified.

CHAPTER 5. RESULTS

This chapter presents the results from the statistical analyses and is divided into three segments. The first section reports the descriptive statistics of the sampled driver population. The second section provides some descriptive statistics and plots using the roadway attribute data and the in-vehicle survey data. The third section provides the results from two estimated logit models, which identified factors associated with changes in individual driver roughness rankings and segment acceptability by respondents. Various statistical software packages were used to perform these analyses, including LIMDEP, SPSS, SST, and SAS.

DESCRIPTIVE STATISTICS OF PARTICIPANTS

Before the data from the in-vehicle survey are presented, it is important and constructive to understand the sample populations that comprised the preliminary survey respondents and eventual pool of driving subjects. The results from simple summary statistics are presented in the following tables and figures.

Figure 8 shows that approximately 60 percent of the respondents were men and 40 percent were women in both the preliminary and in-vehicle surveys. While it would have been desirable to have a split of men and women (approximately 50/50) that was representative of the drivers in the State of Washington, this distribution was acceptable.

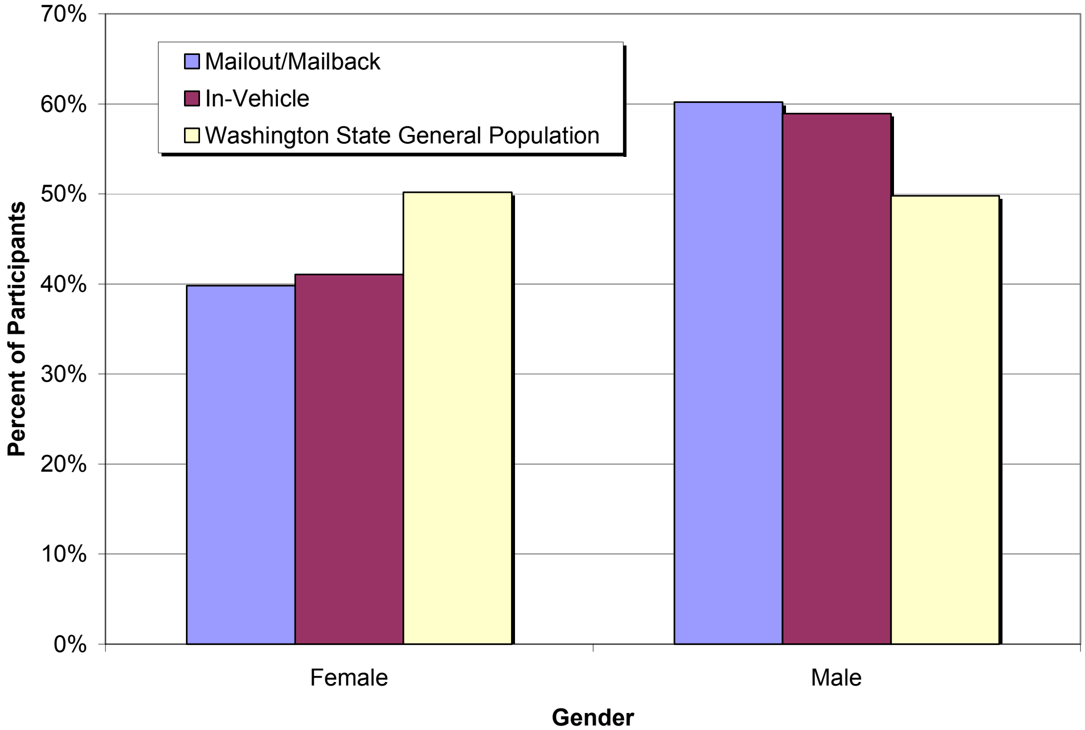
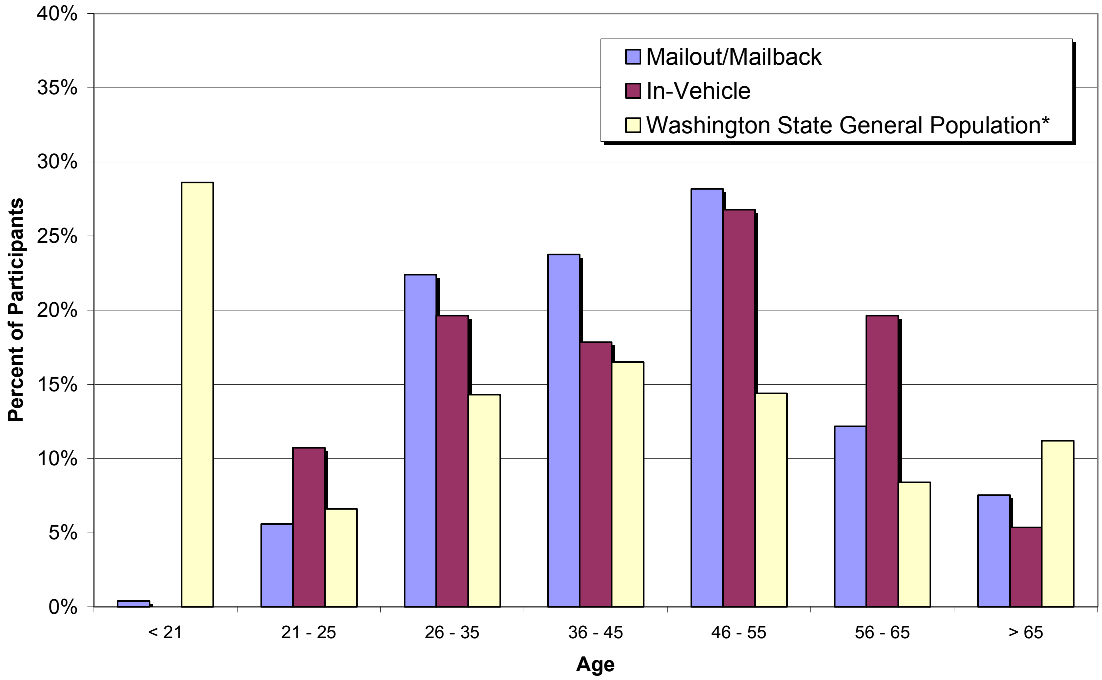


Figure 8. Gender of Participants

As indicated in Figure 9, the age distribution was relatively uniform for most age groups. In general, approximately 10 percent of all age groups were represented by the in-vehicle participants, except for the youngest and oldest age groups. It is possible that these groups were underrepresented because they drive less often than middle-aged individuals. Fewer individuals over the age of 65 participated in the study. Initially surprising was the small number of participants under the age of 21, given the fact that participants were randomly selected near the University of Washington. However, this representation is not so surprising given that many students under 21 may not own a car that is registered in their name and, therefore, would not have received the preliminary survey.



*Approximate. Census population figures are grouped as: < 20, 20 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, and > 65.

Figure 9. Reported Age of Participants

Figure 10 shows that this sample of individuals was highly educated. Nearly 75 percent of the sampled population had a college or post-graduate degree. By contrast, the latest U.S. Census figures reveal that the Seattle area has a well-educated population, in which approximately 35 percent have a bachelor degree or higher, and the Washington State average is approximately 30 percent (US Census, unpublished).[13] The large number of college graduates was certainly attributable, in part, to selecting an initial pool of potential participants from vehicles in the proximity of the University of Washington. The large number of college graduates also represented a potential bias among the survey participants. As will be noted in the last chapter, the characteristics of the drivers, as well

[13] Nationally, approximately 26 percent of individuals have a bachelor degree or higher (Census, unpublished).

as the characteristics of the test segments, may have been too unique to extrapolate these results statewide or nationally.

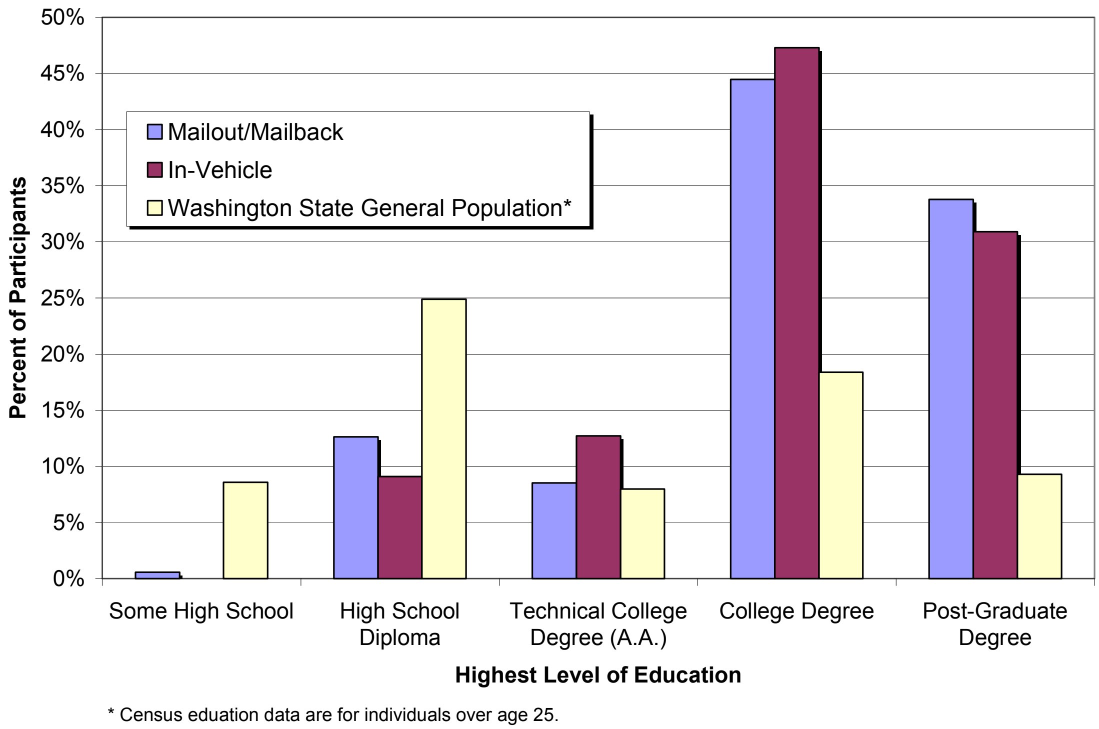


Figure 10. Reported Highest Level of Education of Participants

Along with being highly educated, the sampled population was also financially above average. While a relatively uniform distribution of annual incomes is shown in Figure 11, a large proportion of respondents fell into the two highest income groups, and the lowest income groups were underrepresented. This distribution of higher incomes may have been attributable to the education levels shown in Figure , but it may also have been attributable to the affluent areas around the University of Washington and in urban Seattle/Bellevue.

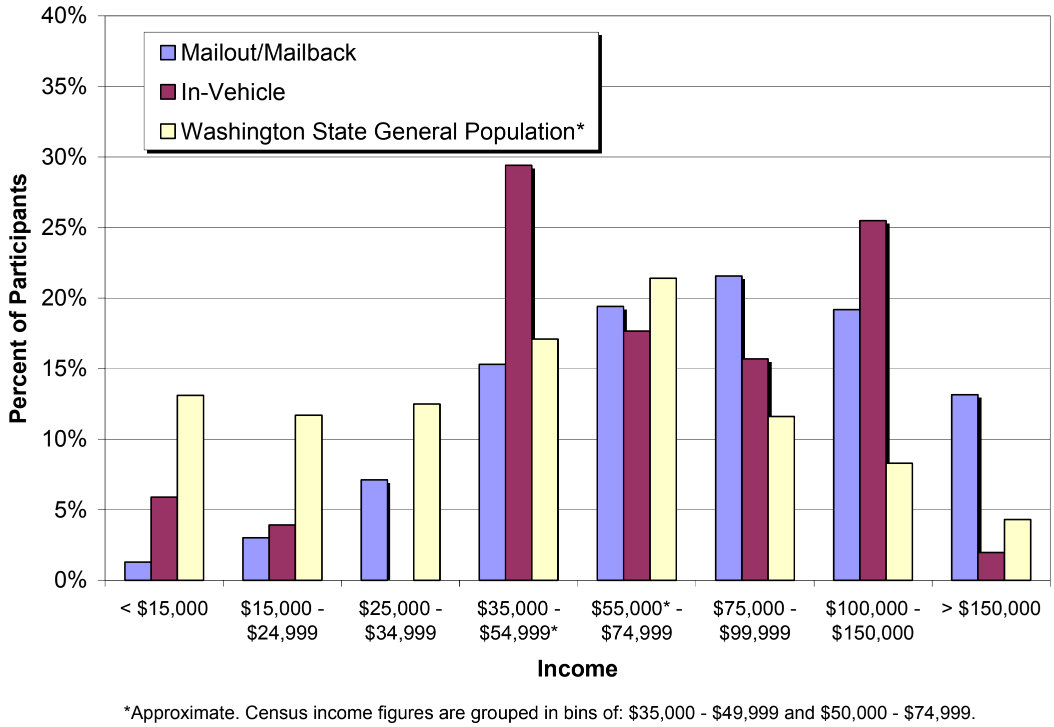


Figure 11. Reported Household Income of Participants

GENERAL PERCEIVED ROUGHNESS FROM PRELIMINARY SURVEY

One of the primary pieces of data received from the preliminary survey provided information on drivers' general perception of road roughness in the Seattle area. Respondents were asked about their general opinion of roughness at various general locations in the area. (See Appendix A.) For example, respondents were asked, "How would you rank the roughness of the highway on I-5 near downtown Seattle?" Respondents were given a choice of 1 (very smooth) through 5 (very rough). (A response of "I do not drive that segment" was also possible.) The results from these data, shown in Table 17, indicate that people had relatively "average" perceptions of roughness around the Seattle area and that there were clear differences in perceived roughness at different locations.

Table 17. Average Perceived Roughness from Preliminary Survey

Location	Mean	Standard Deviation
I-5 near SeaTac Airport	3.05	0.87
I-5 near Downtown Seattle	3.33	0.87
I-5 near Northgate Mall	3.28	0.87
I-90 near Mercer Island	2.15	0.81
I-90 near Bellevue	2.52	0.87
I-90 near Issaquah	2.49	0.93
I-405 near Southcenter	2.97	0.86
I-405 near Bellevue	2.86	0.87
I-405 near Bothell	2.84	0.93
SR520 near Montlake Bridge	3.13	0.88
SR520 near Bellevue	2.66	0.86

Notes: "Perceived roughness" is on a 1 to 5 integer scale, with 1 being "very smooth" and 5 being "very rough". Responses were neglected if the respondent indicated that they do not normally drive on that segment.

While these responses provided useful information, they had limited value for additional analysis because they could not be correlated to specific segments or roadways. The respondents were not provided with specific location information; they were not provided with lane indications or even direction of travel. In other words, respondents were not asked to evaluate segments at a particular location, direction, or lane. This type of information, however, was captured by the in-vehicle survey. The results from those data are presented in the next section.

SUMMARY STATISTICS OF DRIVER ROUGHNESS RANKINGS WITH IRI MEASUREMENTS

After the sample of drivers was reviewed, some of the collected data could be used to generate summary statistics to shed some light on the relationship between drivers' perceptions of roughness and how those perceptions differed from roughness measurements (i.e., IRI).

One of the fundamental questions underlying this research was, How accurately do IRI and physical roughness measurements mimic drivers' perceptions of roughness? Figure 12 is a boxplot that shows the mean IRI value for each roughness ranking. This boxplot indicates, in general, that roughness rankings (1 = "very smooth," 5 = "very rough") increased as the IRI measurements increased, as expected.

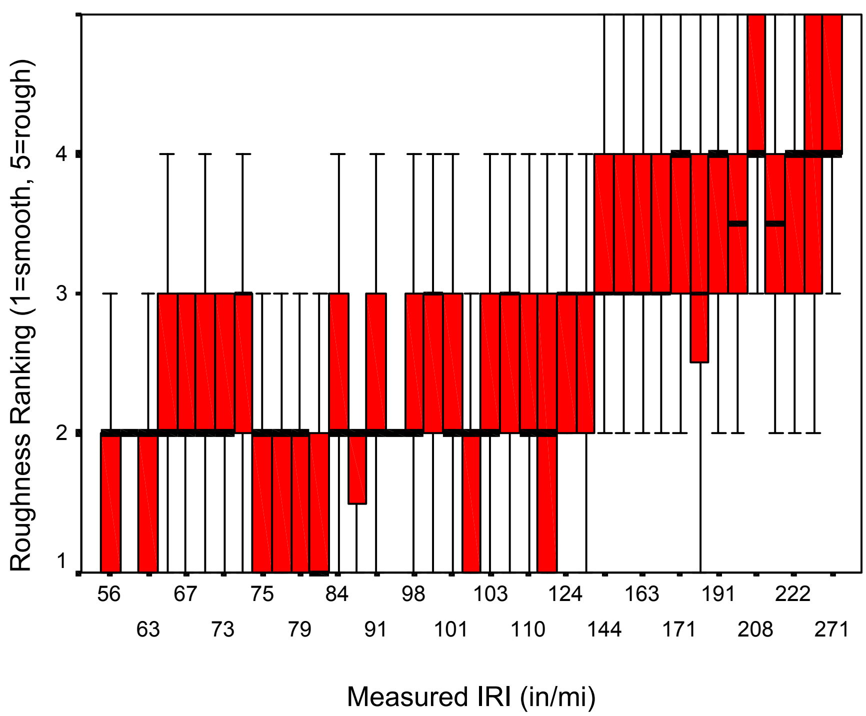


Figure 12. Boxplot of Roughness Ranking vs. Measured IRI

The same result is shown in Figure 13, which is a plot of mean roughness ranking for each test segment and its subsequent IRI value. If the roughness ranking were to be modeled as the dependent variable in a simple-factor linear regression, IRI alone would account for 83 percent of the variability in the average roughness rankings, as denoted by the correlation coefficient, R^2.

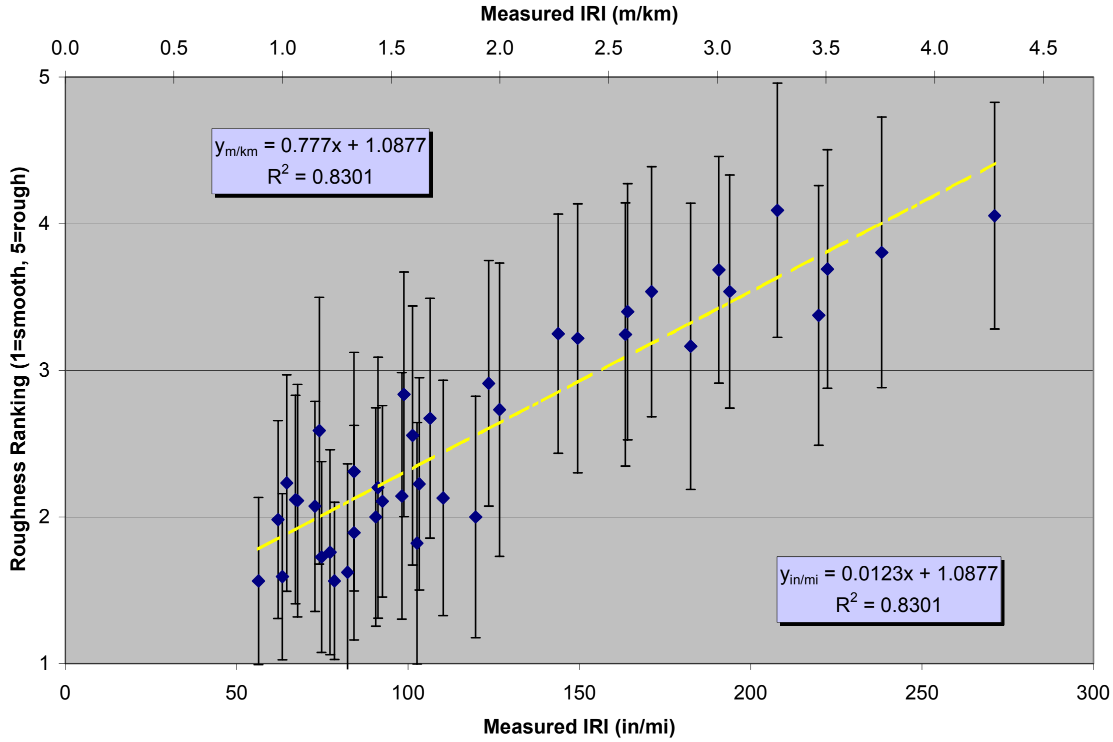


Figure 13. Mean Roughness Ranking with One Standard Deviation vs. Measured IRI

Figure 14 uses the same data as Figure 12, but its axes are reversed. The purpose of this plot is to show how varying IRI measurements corresponded to discrete roughness rankings. This figure, which shows the number of readings, N, in each roughness category, indicates that drivers seemed to be in general agreement about the test segments that they considered to be "smooth," but they had varying opinions about which test segments they considered to be "rough."

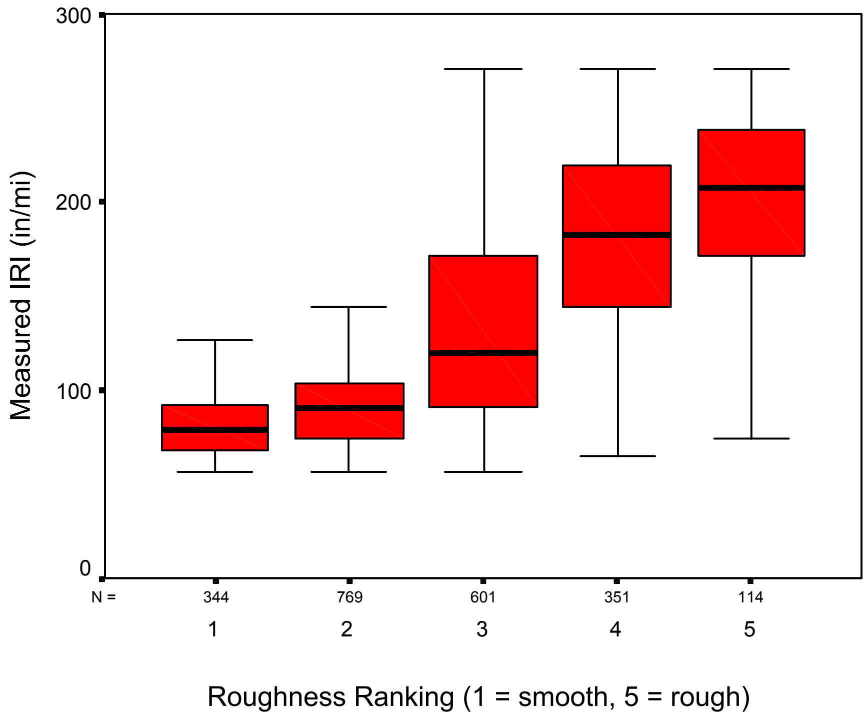


Figure 14. Boxplot of IRI vs. Roughness Ranking

The relationship between roughness rankings (1 = "very smooth," 5 = "very rough") and IRI can be examined further by plotting the average roughness rankings (denoted by the column chart) and IRI measurements (denoted by the line plot) for each roadway of the 40 test segments, as shown in Figure 15. This plot indicates that, while IRI does seem to follow roughness rankings, the two plots are not identical. This finding further supports the idea that driver perception of roughness is probably influenced by factors other than IRI alone.

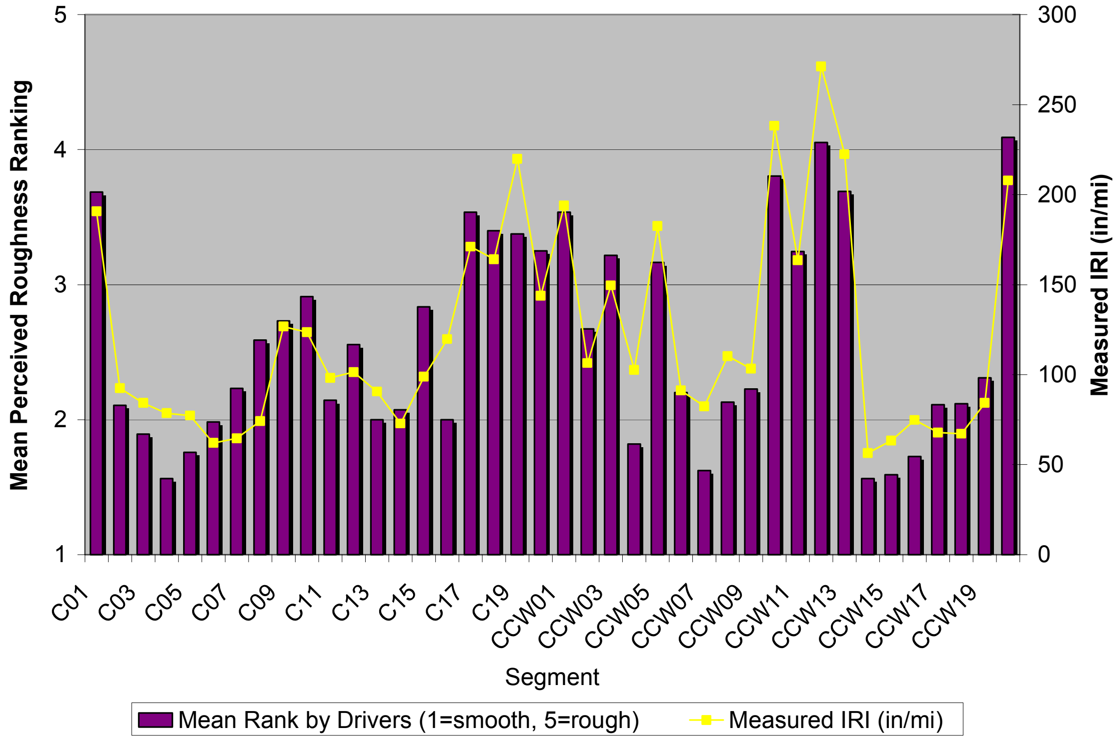


Figure 15. IRI Measurements vs. Driver Roughness Rankings

Figure 16 is a cumulative percentage plot, showing roughness rankings for all sections (sorted by IRI along the x-axis).[14] As expected, the majority of "smooth" rankings were given to test segments with low IRI measurements, and the majority of "rough" test segments were given to test segments with high IRI measurements. As the IRI measurements increase, the rankings generally increase in roughness.

[14] Note that the IRI measurements on the x-axis are not on a linear scale.

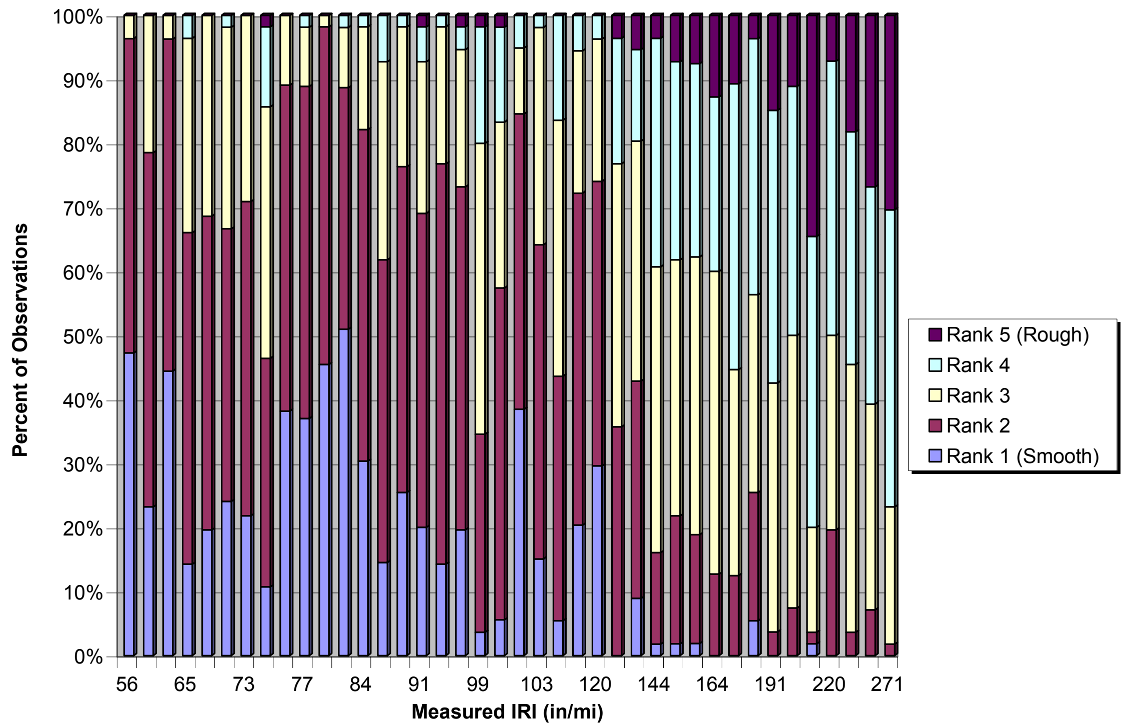


Figure 16. Cumulative Percentage Plot of Roughness Rankings on Each Test Section vs. Corresponding Ordered IRI Measurements

SUMMARY STATISTICS FOR SUBSETS OF DRIVER ROUGHNESS RANKINGS WITH IRI MEASUREMENTS

Next, the data set was divided and plotted in subsets. By plotting different subsets of data, it would be possible to observe unusual trends and to identify factors contributing to changes from the plots shown in the previous section. The roughness rankings were grouped for different sets of individuals (e.g., men vs. women), different test vehicle types, and different roadway attributes (e.g., elevated vs. non-elevated test segments).[15]

Table 18 contains the same data that were used to create Figure 15, but it also provides other information that can be used to create new plots, such as Figure 17 or

[15] It is important to remember that smaller sample sizes accompany the subsets.

Table 18. Summary of Driver Roughness Rankings with Selected Roadway Characteristics

Segment	Mean Roughness Rank (Std. Dev.)	Measured IRI	Surface Type	Elevated
C01	3.69 (0.597)	191 in/mi (3.01 m/km)	PCCP	yes
C02	2.11 (0.425)	93 in/mi (1.46 m/km)	ACP	yes
C03	1.89 (0.534)	84 in/mi (1.33 m/km)	ACP	yes
C04	1.56 (0.288)	79 in/mi (1.24 m/km)	ACP	no
C05	1.76 (0.489)	77 in/mi (1.22 m/km)	ACP	no
C06	1.98 (0.454)	62 in/mi (0.98 m/km)	ACP	no
C07	2.23 (0.545)	65 in/mi (1.02 m/km)	ACP	no
C08	2.59 (0.828)	74 in/mi (1.17 m/km)	ACP	no
C09	2.73 (1.000)	127 in/mi (2.00 m/km)	ACP	no
C10	2.91 (0.701)	124 in/mi (1.95 m/km)	PCCP	no
C11	2.14 (0.706)	98 in/mi (1.55 m/km)	PCCP	no
C12	2.56 (0.780)	101 in/mi (1.60 m/km)	PCCP	no
C13	2.00 (0.556)	91 in/mi (1.43 m/km)	PCCP	no
C14	2.07 (0.513)	73 in/mi (1.15 m/km)	PCCP	yes
C15	2.84 (0.695)	99 in/mi (1.56 m/km)	PCCP	no
C16	2.00 (0.679)	120 in/mi (1.89 m/km)	PCCP	no
C17	3.54 (0.726)	171 in/mi (2.70 m/km)	PCCP	yes
C18	3.40 (0.763)	164 in/mi (2.59 m/km)	PCCP	no
C19	3.38 (0.784)	220 in/mi (3.47 m/km)	PCCP	yes
C20	3.25 (0.664)	144 in/mi (2.27 m/km)	PCCP	yes
CCW01	3.54 (0.631)	194 in/mi (3.06 m/km)	PCCP	yes
CCW02	2.67 (0.669)	106 in/mi (1.68 m/km)	PCCP	yes
CCW03	3.22 (0.840)	150 in/mi (2.36 m/km)	PCCP	no
CCW04	1.82 (0.678)	103 in/mi (1.62 m/km)	PCCP	no
CCW05	3.16 (0.954)	182 in/mi (2.88 m/km)	PCCP	no
CCW06	2.20 (0.793)	91 in/mi (1.44 m/km)	PCCP	yes
CCW07	1.62 (0.547)	82 in/mi (1.30 m/km)	PCCP	no
CCW08	2.13 (0.643)	110 in/mi (1.74 m/km)	PCCP	no
CCW09	2.23 (0.545)	103 in/mi (1.63 m/km)	PCCP	no
CCW10	3.80 (0.852)	238 in/mi (3.76 m/km)	PCCP	no
CCW11	3.25 (0.804)	163 in/mi (2.58 m/km)	PCCP	no
CCW12	4.05 (0.597)	271 in/mi (4.28 m/km)	PCCP	no
CCW13	3.69 (0.662)	222 in/mi (3.51 m/km)	PCCP	no
CCW14	1.56 (0.325)	56 in/mi (0.89 m/km)	ACP	no
CCW15	1.59 (0.322)	63 in/mi (1.00 m/km)	ACP	no
CCW16	1.73 (0.424)	75 in/mi (1.18 m/km)	ACP	no
CCW17	2.11 (0.629)	68 in/mi (1.07 m/km)	ACP	yes
CCW18	2.12 (0.506)	67 in/mi (1.06 m/km)	ACP	yes
CCW19	2.31 (0.662)	84 in/mi (1.33 m/km)	ACP	yes
CCW20	4.09 (0.751)	208 in/mi (3.28 m/km)	PCCP	yes

Note: Roughness ranking is on a 1 to 5 integer scale, with 1 being "very smooth" and 5 being "very rough."

Figure 18. Figure 17 identifies some test segments as being on an elevated structure, such as a bridge. While some bridge segments were ranked as being rougher than the IRI alone would suggest, the results are not conclusive. It remains possible that bridge segments are perceived by drivers to be rougher, in general, than non-bridge segments.

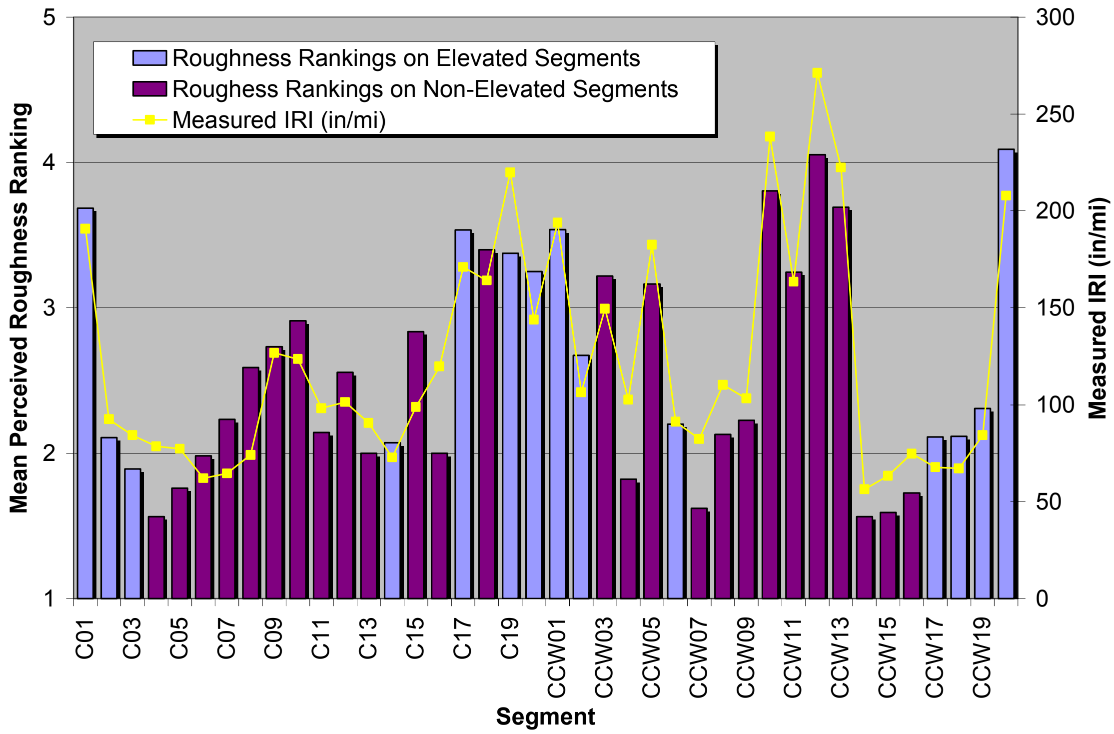


Figure 17. IRI Measurements and Average Roughness Rankings on Elevated and Non-Elevated Segments

Correspondingly, few conclusions can be drawn from Figure 18. It is not clear from this figure whether drivers ranked one pavement type as being rougher than the other. Because this study had a limited sample of pavement sections, it remains possible that any trends could be attributable to other factors.

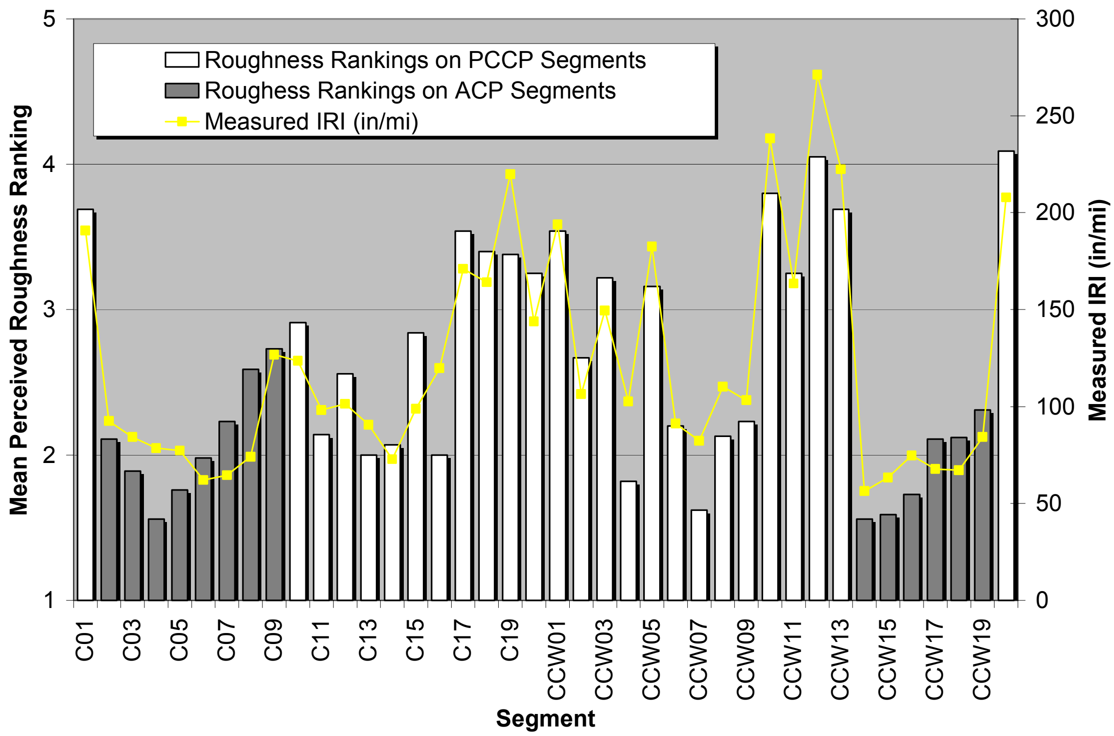


Figure 18. IRI Measurements and Average Roughness Rankings by Surface Type

In general, it is difficult to draw conclusions from figures like 17 or 18. While one could study individual segments and speculate why the roughness rankings and IRI deviate, it is important to realize that the subjective driver roughness rankings could have been influenced by attributes other than the fact that the segment was on an elevated or non-elevated cement concrete or asphalt segment.

The next subset of data analyzed was test vehicle type. The mean roughness rankings for all four test vehicle types are presented in Figure 19, but they do not indicate any clear differences. If cumulative percentage plots are created for each vehicle type, it might be possible to see differences in roughness rankings that might not be apparent by looking at the means in Figure 19. The cumulative percentage plots for each vehicle type are shown in Figure 20 through FIGURE 23. A comparison of these plots could indicate a tendency for drivers to perceive less roughness (i.e., fewer 5 rankings) in sport-utility vehicles (SUVs) relative to other vehicle types, as shown in Figure 2. Additionally, there was a tendency for drivers in minivans to perceive more smoothness (i.e., more 1 and 2 rankings) relative to other vehicle types, as shown in FIGURE 23. With all of the cumulative plots, it is important to remember that the sample size for each vehicle type was approximately one-quarter of the total sample size. As a result, it is possible that the observations within each vehicle type were an artifact of the smaller sample sizes.

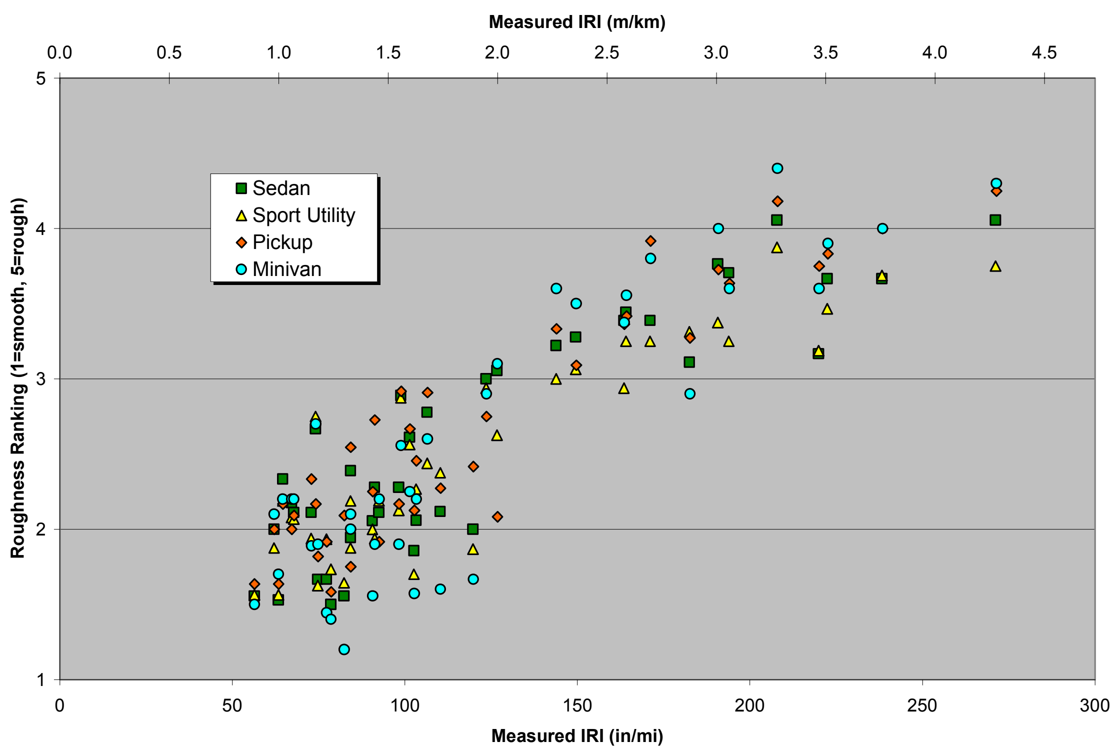


Figure 19. Roughness Ranking vs. Measured IRI by Vehicle Type

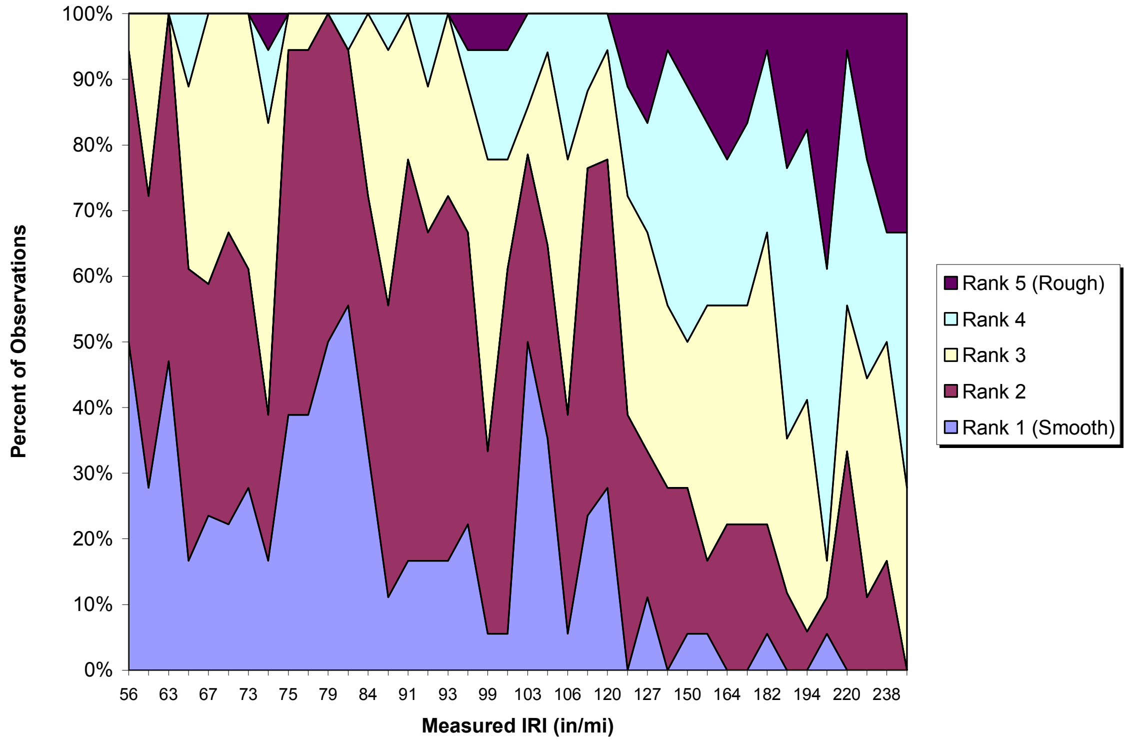


Figure 20. Roughness Ranking vs. Measured IRI for Subset of Drivers in Midsize Sedan Test Vehicle

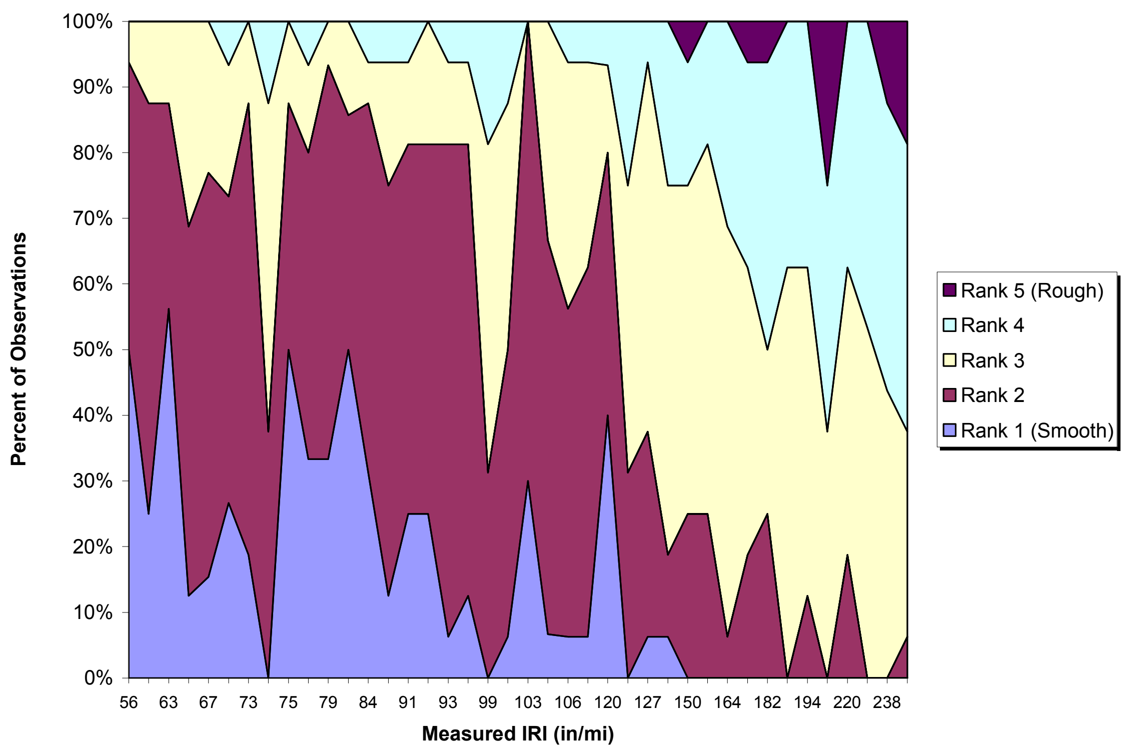


Figure 21. Roughness Ranking vs. Measured IRI for Subset of Drivers in Sports Utility Test Vehicle

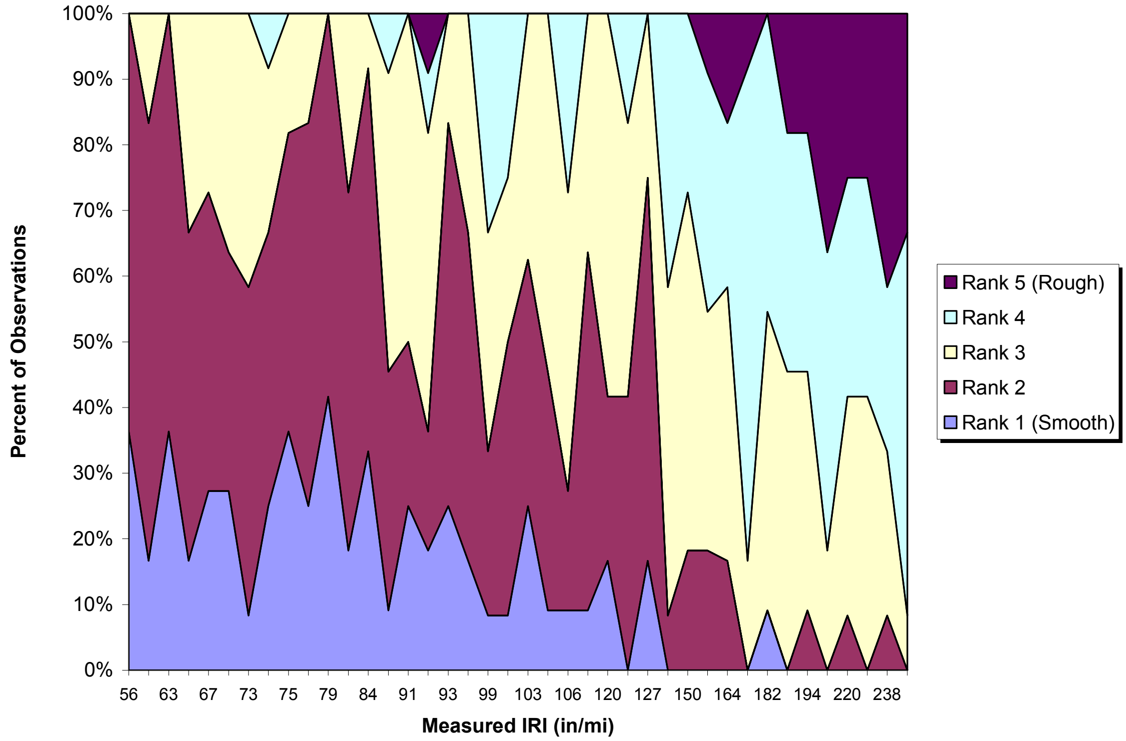


Figure 22. Roughness Ranking vs. Measured IRI for Subset of Drivers in Pickup Truck Test Vehicle

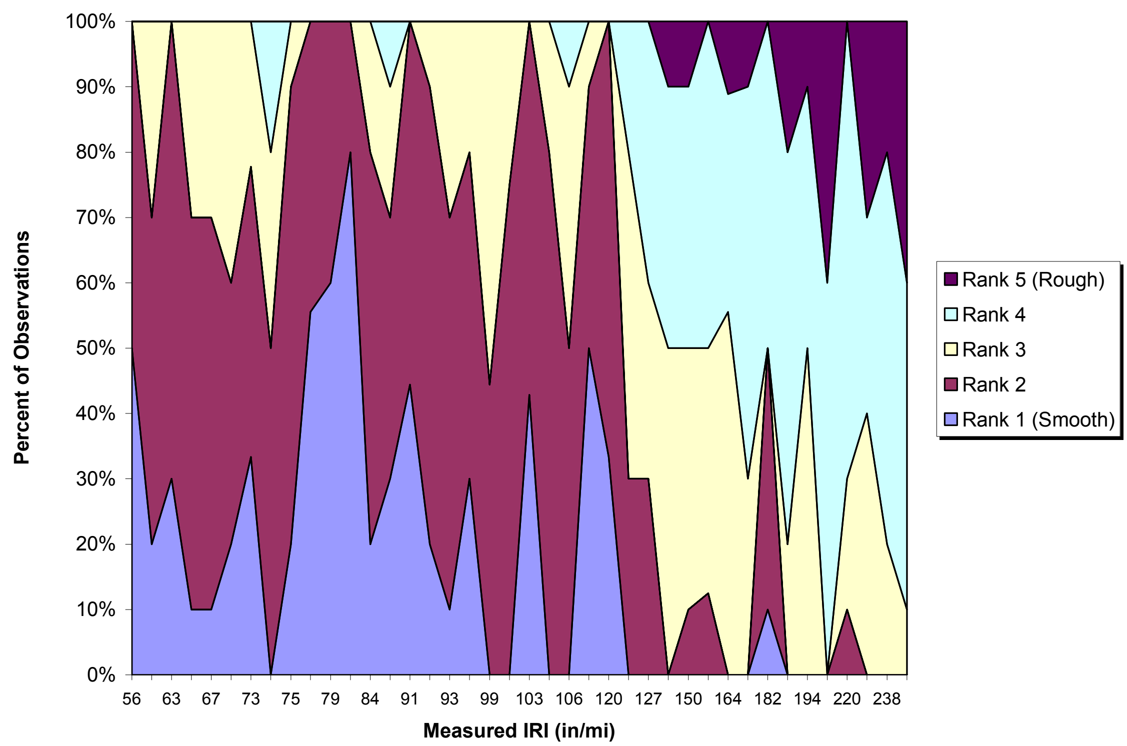


Figure 23. Roughness Ranking vs. Measured IRI for Subset of Drivers in Minivan Test Vehicle

A similar analysis of individual attributes, such as gender, reveals little by visual inspection. A comparison of mean values is shown in Figure 24, and cumulative percentage plots of roughness rankings by female and male drivers are shown in Figure 25 and Figure 26, respectively.

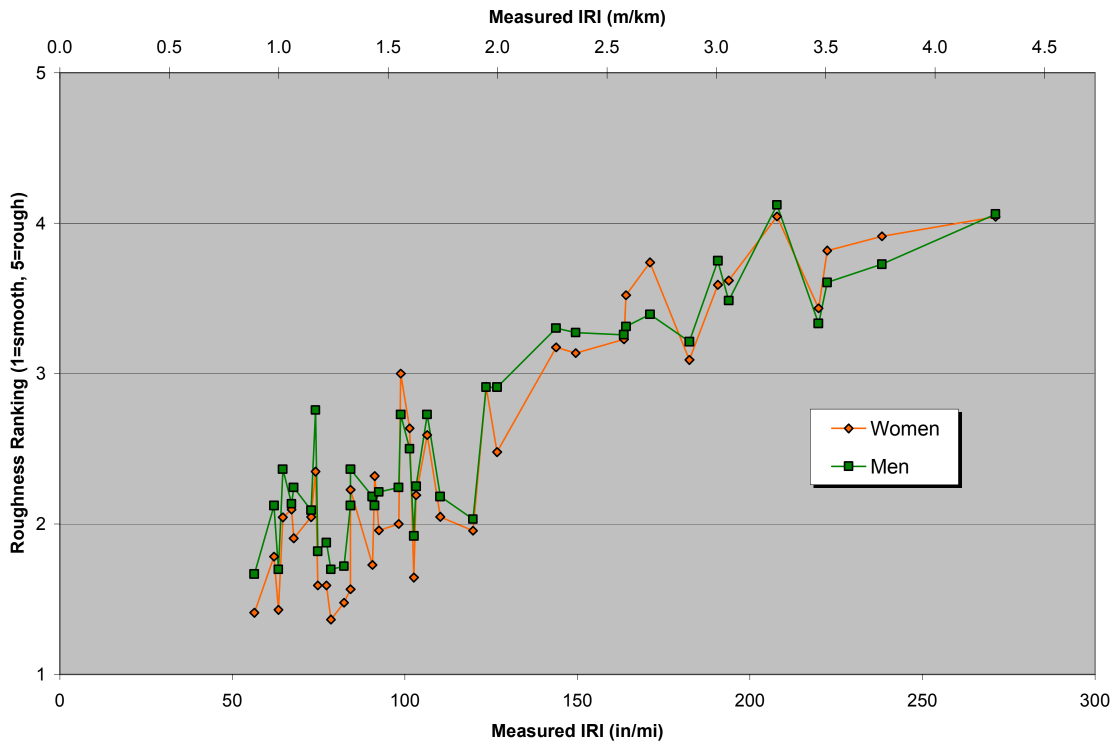


Figure 24. Roughness Ranking vs. Measured IRI by Gender

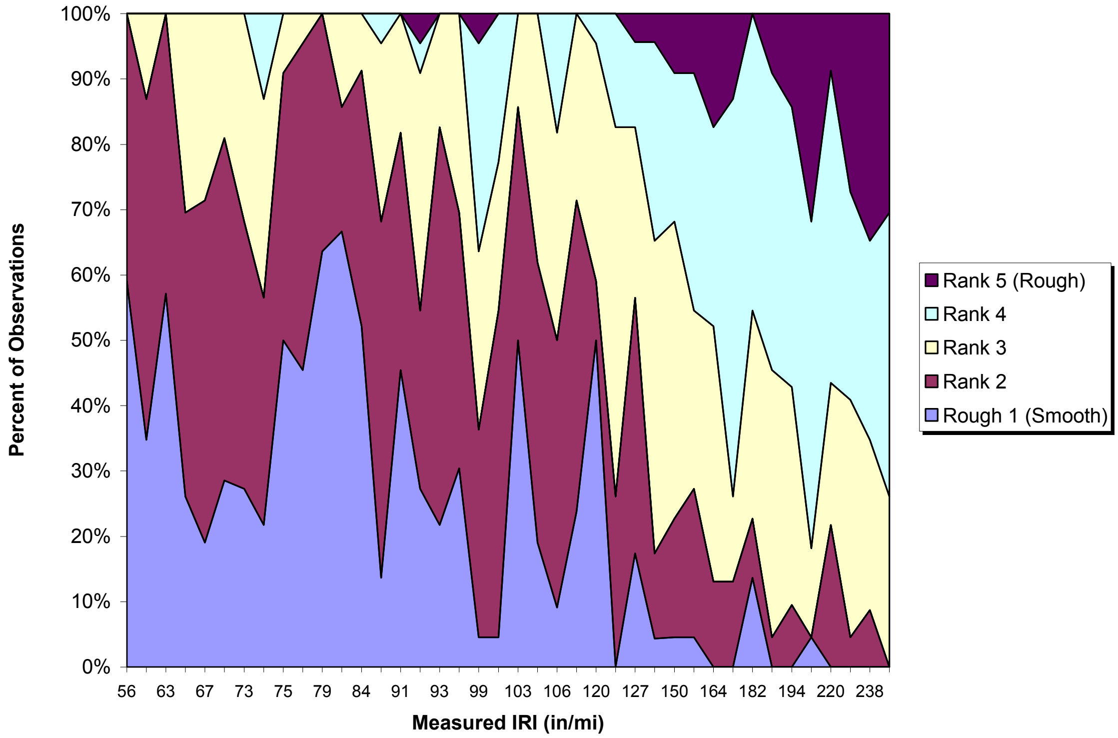


Figure 25. Roughness Ranking vs. Measured IRI for Subset of Female Drivers

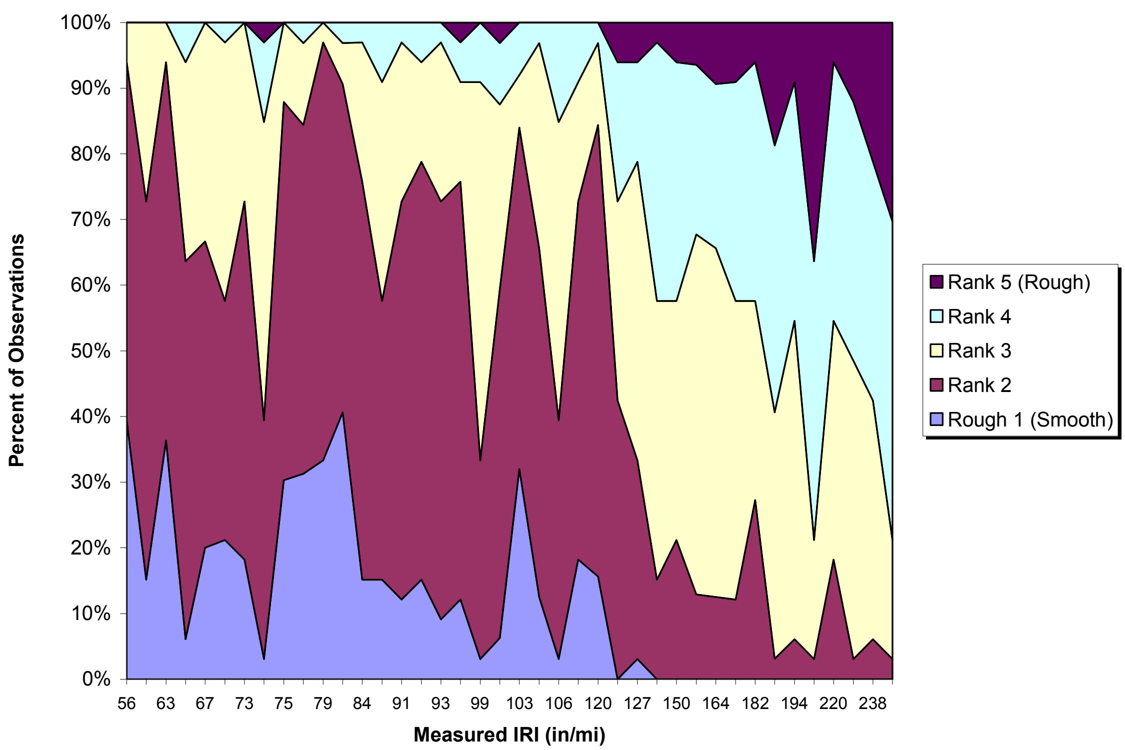


Figure 26. Roughness Ranking vs. Measured IRI for Subset of Male Drivers

While a visual inspection of trends in subsets of the data are helpful at the exploratory stage and can be easy to interpret, it would be more efficient to identify statistically significant contributions to the variance in the perceived roughness rankings through various statistical tests. While the reader is encouraged to make his or her own conclusions from charts such as these, caution should be taken, particularly with the comparison of roughness rankings and IRI. Roughness rankings are not measured with the same precision as IRI, and they are not measured on the same scale.[16] These figures are helpful in recognizing general trends, but caution should be exercised when focusing on one or two specific segments that may have other unusual characteristics. Statistical methods complement these exploratory figures and provide a more conclusive link between roughness rankings and individual, roadway, or vehicle attributes, as explored in the last section of this chapter.

RELATIONSHIP BETWEEN DRIVER ROUGHNESS RANKINGS AND DRIVER BEHAVIOR

Another interesting relationship that was explored with these data is the relationship between roughness rankings and driver behavior variables, specifically driving speed. When a boxplot is made of speed versus roughness ranking and speed outliers (due to heavy congestion) are removed, Figure 27 is obtained. It indicates that there is little impact on driving behavior from differences in perceived road roughness at these roughness levels.

---

[16] IRI has a lower bound of zero and does not have an upper bound, while roughness rankings have a confined range of one to five.

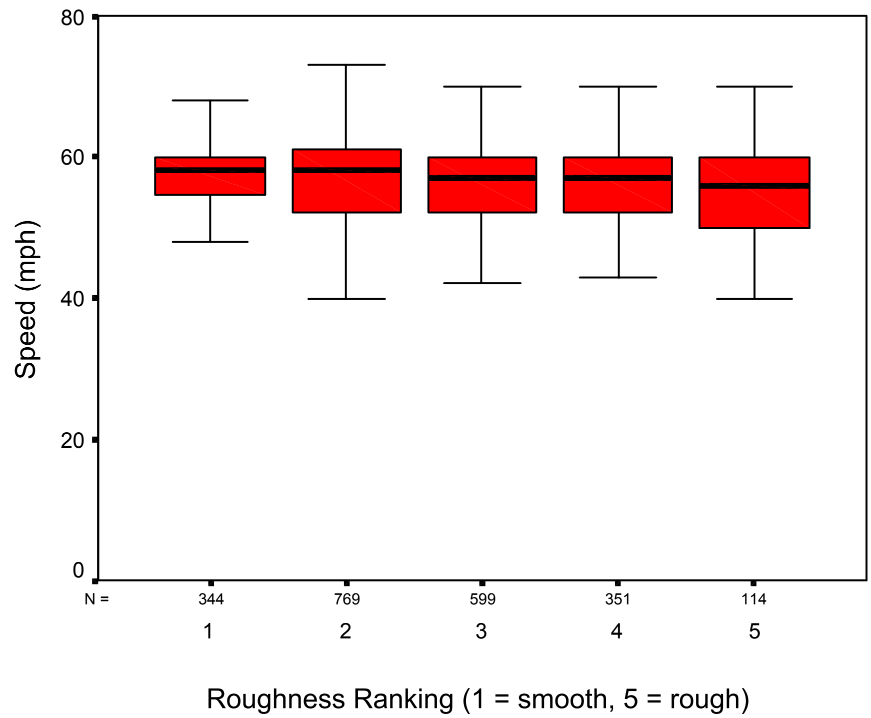


Figure 27. Boxplot of Speed vs. Roughness Rankings

POTENTIAL ACCEPTABLE/UNACCEPTABLE IRI THRESHOLDS

Another goal of this research was to identify potential roughness thresholds at which the majority of drivers found the roughness levels to be distinctly acceptable or unacceptable. These threshold values are important because the point that drivers agree to be distinctly acceptable could be a specification limit during the construction phase, and the point that drivers agree to be distinctly unacceptable could be an indicator for necessary rehabilitation.

Tables 19 and 20 are organized according to WSDOT specifications for ranking pavement condition according to IRI. WSDOT uses the following five categories:

Very Good	<95 in/mi.
Good	95-170 in/mi.

Fair	170-220 in/mi.
Poor	220-320 in/mi.
Very Poor	>320 in/mi.

When a pavement is classified to be in the Fair condition, this correlates to a pavement requiring rehabilitation, which for WSDOT is typically a 45mm (1.8 inch) asphalt concrete overlay.

First, it was helpful to sort all observations by IRI and group them into acceptable and unacceptable bins, as in Table 19. This table clearly shows that the acceptability ratings decreased as the roughness increased. The fraction of acceptability ratings dropped most dramatically when IRI increased from range 2, at which 83 percent of the ratings were considered to be acceptable, to range 3, at which only 60 percent of the ratings were considered to be acceptable. The change in IRI ranges resulted in a drop of 26 percent in acceptability ratings.

Table 19. Distribution of Acceptable/Unacceptable Observations, Grouped by IRI

IRI Grouping	Observations	Acceptable Rating	Unacceptable Rating
< 95 in/mi (< 1.5 m/km)	931	905 (97.2%)	26 (2.8%)
95 – 170 in/mi (1.5 – 2.7 m/km)	751	626 (83.4%)	125 (16.6%)
170 – 220 in/mi (2.7 – 3.5 m/km)	331	197 (59.5%)	134 (40.5%)
220 – 320 in/mi (3.5 – 5.0 m/km)	167	81 (48.5%)	86 (51.5%)
≥ 320 in/mi (≥ 5.0 m/km)	N/A	N/A	N/A

Note: The acceptable and unacceptable ratings for each IRI grouping total 100 percent. None of the pavement sections included in the study had an IRI ≥ 320 in/mi.

Table 20 has identical information but groups all observations into cumulative IRI groupings. In this format, the change in driver perception is less apparent. This table is useful because it illustrates that drivers, on a whole, seemed relatively pleased with the roads in the study area. Over the entire sample, only 17 percent of the observations were still deemed to be unacceptable, as shown in the last 4th IRI range of Table 20. It is important to recognize that the sample of test segments was not uniformly distributed and that it consisted primarily (82 percent) of test segments with an IRI of less than 190 in/mi (3.0 m/km). The fact that the sample was more heavily weighted with smooth roads could also have contributed, in part, to the high number of acceptable evaluations.

Table 20. Cumulative Distribution of Acceptable/Unacceptable Observations, Grouped by IRI

Cumulative IRI Grouping	Cumulative Observations	Acceptable Rating	Unacceptable Rating
< 95 in/mi (< 1.5 m/km)	931	905 (97.2%)	26 (2.8%)
95-170 in/mi (1.5-2.7 m/km)	1682	1531 (91.0%)	151 (9.0%)
170-220 in/mi (2.7-3.5 m/km)	2013	1728 (85.8%)	285 (14.2%)
220-320 in/mi (3.5-5.0 m/km)	2180	1809 (83.0%)	371 (17.0%)
≥320 in/mi (≥ 5.0 m/km)	N/A	N/A	N/A

Note: The acceptable and unacceptable ratings for each IRI grouping total 100 percent. None of the pavement sections included in the study had an IRI ≥ 320 in/mi.

If the proportion of acceptable and unacceptable ratings *for each segment* is plotted against IRI, Figure 28 is obtained.[17] For segments with low IRI measurements

[17] There is an acceptable percentage and an unacceptable percentage for each segment on the x-axis. In reality, only one point (i.e., the percentage of acceptable ratings) for each segment is necessary because the other point (i.e., the percentage of unacceptable ratings) could be obtained by subtracting one from 100 percent to obtain the other.

(e.g., less than 100 in/mi or 1.6 m/km), nearly all ratings were “acceptable,” and only a fraction were “unacceptable.” For test segments that exceeded the FHWA IRI recommendation, there was little agreement about what was “acceptable” and what was “unacceptable.” This figure shows that a pavement section at the FHWA IRI recommendation of 170 in/mi (2.7 m/km) would be expected to have around a 65 percent acceptability rating (a 35 percent unacceptable rating).

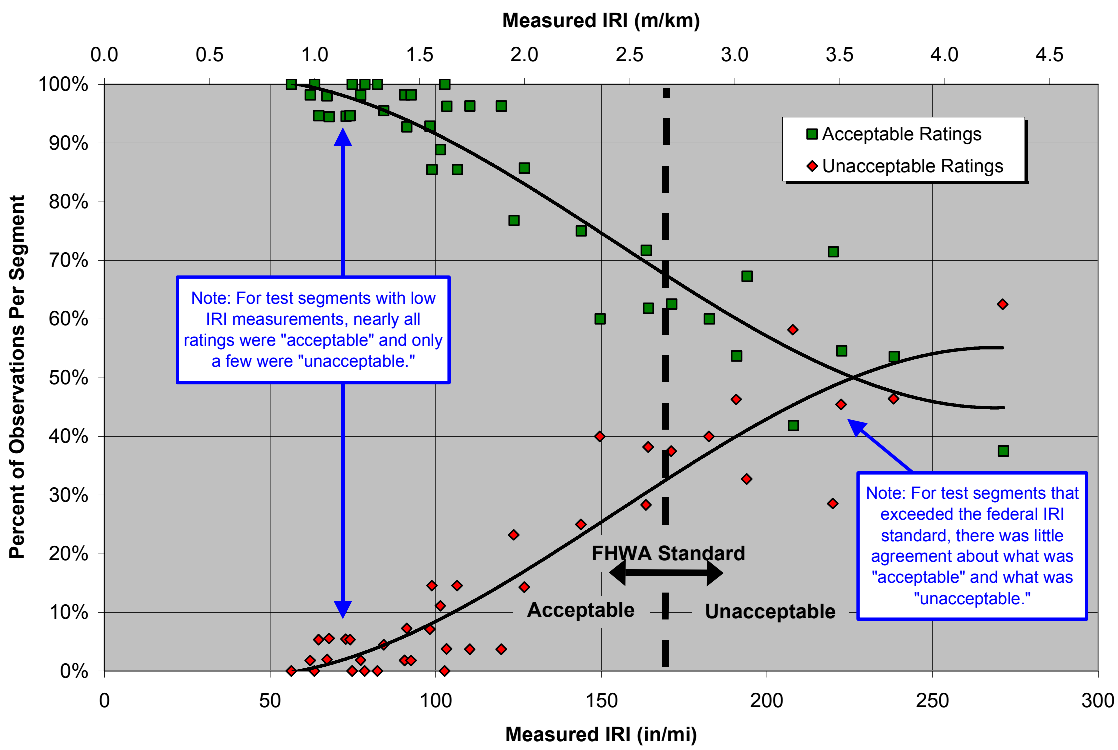


Figure 28. Proportion of Acceptable and Unacceptable Ratings for Each Test Segment (Sorted by IRI)

The acceptable and unacceptable data can also be inspected as collective groups. Figure 29 shows the cumulative percentage of all of the nearly 1,800 acceptable ratings against the IRI values on those segments. In other words, this figure contains all acceptable ratings and indicates how the collective total is obtained as IRI increases.

Nearly half of all “acceptable” ratings were made on segments with an IRI of less than around 90 in/mi (1.42 m/km). This figure also indicates that about 90 percent of all “acceptable” ratings were made on test segments with an IRI of 170 in/mi (2.7 m/km) or less.

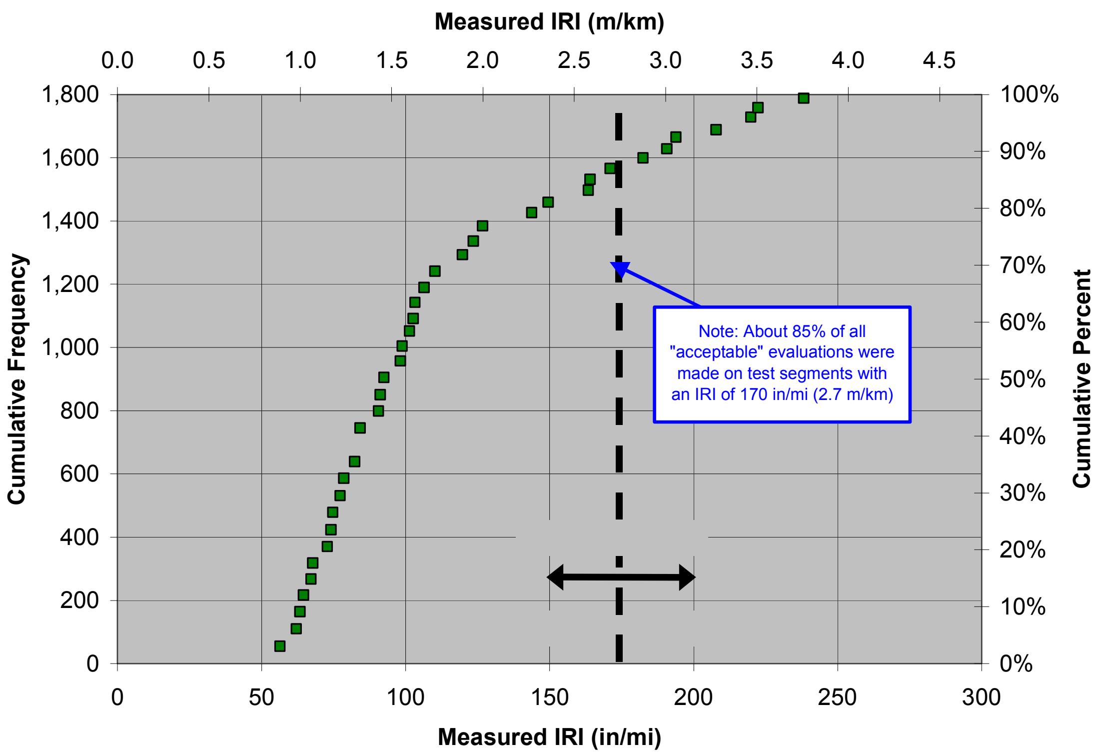


Figure 29. Cumulative Distribution of Acceptable Ratings, Sorted by IRI

Instead of a cumulative distribution, it may be more intuitive to characterize this process as a survival process in which the accumulation of all acceptable ratings starts at 100 percent and eventually decreases to zero, as shown in Figure 30. (Note that the survival function in Figure 30 is simply the cumulative distribution function in Figure flipped over. The survival percentage is simply 100 percent from the cumulative percentage.) In other words, almost all acceptable ratings are made at extremely low IRI levels but the proportion of acceptable ratings decreases as IRI increases. The survival function indicates that only about 10 percent of all acceptable ratings were above the FHWA recommendation.

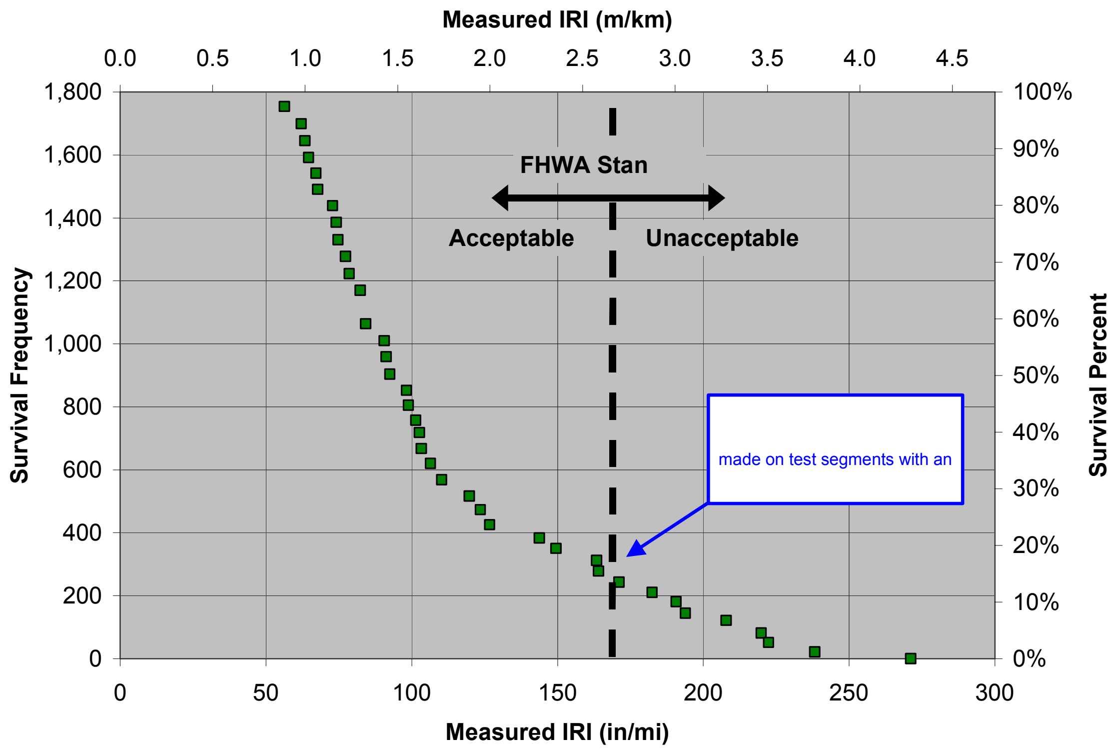


Figure 30. Survival Function of Acceptable Ratings, Sorted by IRI

The same cumulative analysis can be done with the group of unacceptable ratings. Figure 31 is a cumulative percentage plot of the nearly 400 unacceptable ratings against the IRI values on those segments. The group of unacceptable ratings reveals some notable distinctions from the group of acceptable ratings. First of all, about 50 percent of unacceptable ratings still occurred at the FHWA roughness recommendation, even though nearly 90 percent of acceptable ratings were made at that IRI level. This finding indicates that all unacceptable ratings are not made exclusively at high IRI levels and that other factors may influence the fewer number of unacceptability ratings.

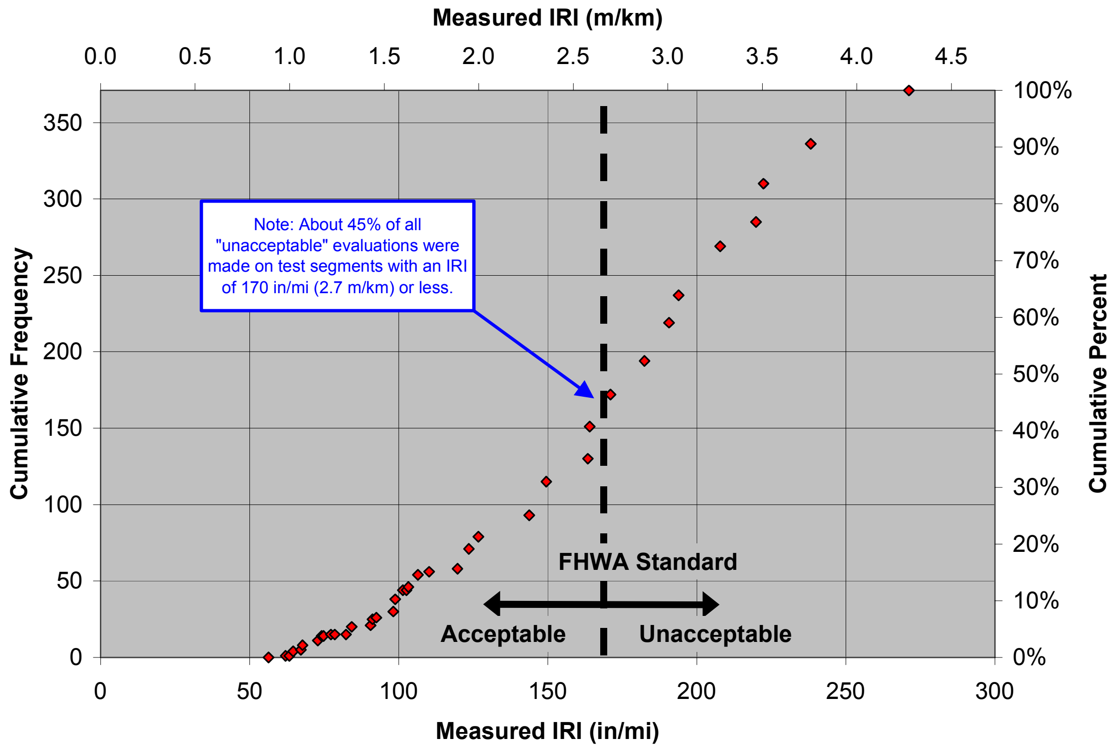


Figure 31. Cumulative Distribution Function of Unacceptable Ratings, Sorted by IRI

When both cumulative distributions (Figure 30 and Figure 31) are compared, additional observations can be made. The first focuses on the "tail" ends of the functions. About 10 percent (190 ratings) of the ratings in Figure 30 were ranked as being acceptable, despite having a corresponding IRI of above 190 in/mi (3.0 m/km). Interestingly, almost 10 percent of the ratings (35 ratings) were "unacceptable" where the IRI was less than 95.0 in/mi (1.5 m/km), as shown in Figure 31.[18] Again, these ratings could support the contention that issues other than IRI affect drivers' perceptions of roughness.

Additional insights can be gained by looking at the histogram of acceptable and unacceptable ratings against IRI. Figure 32 shows the histograms of acceptable and

[18] It is important to note that the two functions in Figure 31 are based on different sample sizes. There were over 1,800 acceptable ratings but fewer than 400 unacceptable ratings.

unacceptable ratings along with their corresponding normal curves. The test segments used in this study had IRI values that were not uniformly distributed but skewed toward low values, and this is also seen in the histograms of acceptable and unacceptable ratings. The normal curves for acceptable and unacceptable ratings crosses over at an IRI value of about 200 in/mi (3.2 m/km), i.e., the number of unacceptable ratings exceeds the number of acceptable ratings for IRI values of 200 in/mi (3.2 m/km) or higher. Alternatively, it indicates that majority of the users will give acceptable ratings for pavement sections that have an IRI value of 200 in/mi (3.2 m/km) or better. Furthermore, it can also be seen that the users will give acceptable ratings by a margin of 2 to 1 or higher for pavement sections that have the FHWA recommended IRI value of 170 in/mi (2.7 m/km) or better.

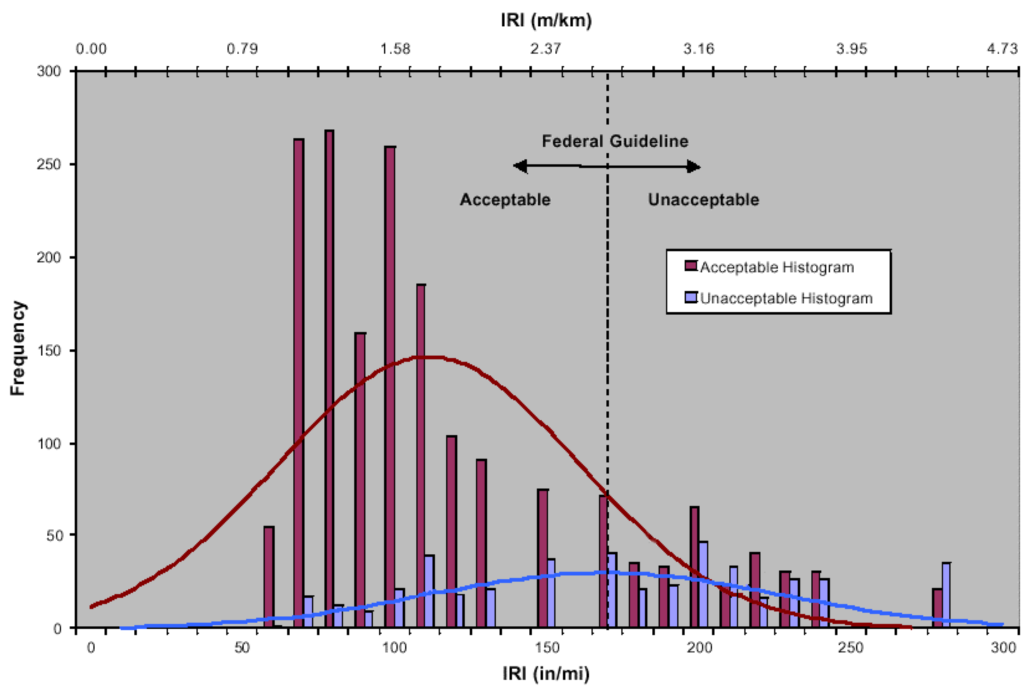


Figure 32. Acceptability and Unacceptability Histograms With Fitted Normal Curves

The difficult question is deciding where the threshold should be placed. In other words, what fraction of the driving population do the policymakers want to satisfy? That is a subjective or economic decision, not an engineering one. After all, it may be unrealistic to satisfy all drivers' tastes and preferences, but it may be useful to identify the values that satisfy a large portion of the driving population, at a reasonable cost. These data can also be tabulate, shown in Table 21, to identify other potential threshold values.

Table 21. Potential IRI Acceptability/Unacceptability Thresholds

Cumulative Percentile	Acceptable IRI		Unacceptable IRI	
	m/km	In/mi	M/km	in/mi
5%	0.98	62	1.33	84
10%	1.00	63	1.56	99
15%	1.06	67	1.73	110
20%	1.14	72	1.97	125
25%	1.18	75	2.27	144
30%	1.22	77	2.34	148
35%	1.29	82	2.58	163
40%	1.32	84	2.59	164
45%	1.43	91	2.67	169
50%	1.46	93	2.81	178
55%	1.56	99	2.93	186
60%	1.62	103	3.02	191
65%	1.68	106	3.09	196
70%	1.82	115	3.22	204
75%	1.97	125	3.39	215
80%	2.36	150	3.49	221
85%	2.61	165	3.56	226
90%	3.01	191	3.76	238
95%	3.47	220	4.00	253
100%	4.28	271	4.28	271

Notes: Some acceptable/unacceptable IRI values were interpolated from IRI values at percentiles above and below the desired percentile. The highest IRI exposed to participants was 271 in/mi (4.28 m/km); the lowest IRI was 62 in/mi (0.89 m/km).

ORDERED LOGIT MODEL OF DRIVER ROUGHNESS RANKINGS

To extend the exploratory analysis and identify other factors associated with changes in driver roughness rankings, econometric models were used. As noted in the previous chapter, the data lent themselves well to an ordered logit model.

The dependent variable was the drivers' roughness rankings, an ordinal variable for which a value of 1 represented smooth ("excellent") pavement quality and a value of 5 represented rough ("poor") pavement quality. This model was appropriate because it can help identify common characteristics that are associated with above- or below-average rankings. In other words, this model would help identify roadway, vehicle, or driver characteristics that are associated with "rougher" or "smoother" than average rankings.

Along with the identification of covariates, this model could provide some indication as to a variable's relative importance on the roughness ranking outcomes and should help prioritize the factors in terms of their associated influence on the individual driver roughness rankings. All of these results could be useful in helping departments of transportation identify the factors that are most associated with perceived "rough" and "smooth" roadways. With this information, they can design and maintain roads that are more suited to the desires of drivers. More importantly, they can prioritize rehabilitation so that it is more in accordance with the driving population.[19]

The results from an ordered logit model are presented in Table 22.

This model provided information about the individual, the vehicle, and the roadway and how they are associated iwth perceived roughness rankings. The t-statistic indicates the significance of the explanatory variable and justifies its inclusion in the model. Typically, a t-statistic value of 1.96 (which corresponds to a 95 percent confi-

[19] This study is not recommending that rehabilitation project prioritization be based on these results. Project prioritization is dependent on many factors, and it would not be economically feasible to base project selection solely on driver preferences. The recommendation is that these factors be used, at the very least, to help prioritize projects if all other issues are equal.

Table 22. Ordered Logit Model of Driver-Perceived Roughness Rankings

Independent Variable	Estimated Coefficient	t-Statistic
Constant	-10.0600	-6.606
Pavement-Specific Variables		
IRI Measurement (in/mi) of roadway test segment	0.0147	9.533
Patch Indicator (1 if the segment appeared to have patch work, 0 otherwise)	0.8670	5.703
Joint/Abutment Indicator (1 if the segment contains expansion joints or bridge abutments, 0 otherwise)	0.5162	4.209
Age of Roadway Segment Surface (years)	0.0163	3.458
Drive-Specific Variables		
SUV or Minivan Test Vehicle Indicator (1 if sport utility or minivan was test vehicle type, 0 otherwise)	-0.7664	-7.197
Noise (dB) inside test vehicle during evaluation	0.1294	6.708
Speed (mph) of test vehicle during evaluation	-0.0196	-4.286
Noise Increase Indicator (1 if the noise inside test vehicle during evaluation increases by 3 dB or more between two adjacent test segments, 0 otherwise)	0.9563	3.879
Interaction between Speed (mph) and measured (PSC)	0.0006	7.260
Individual-Specific Variables		
SR520 User Indicator (1 if participant uses SR520 “always” or “often,” 0 otherwise)	-0.8536	-8.319
I-405 User Indicator (1 if participant uses I-405 “always” or “often,” 0 otherwise)	0.6864	6.944
Gender Indicator (1 if participant was female, 0 otherwise)	-0.5536	-5.538
Older Age Indicator (1 if participant was over age 55, 0 otherwise)	-0.5945	-5.240
Per Capita Household Income	0.1484	4.647
Number of Observations	1936	
Initial Log-Likelihood	-2,836.547	
Log- Likelihood at Convergence	-2,196.879	
Chi-Squared Statistic	1,279.335	
Percentage Correctly Predicted	47.521	

Notes: Responses for the dependent variable are whole number ranks between 1 (very smooth) and 5 (very rough). A positive coefficient indicates that the variable is associated with increased perceived roughness by the respondent.

dence level) is used as a threshold value to determine whether a variable should be included in the model. In this study, higher t-statistics (arbitrarily set at 3.29, which typically corresponds to a 99.9 percent confidence level) were required to account for possible correlation issues, which are explained later in the chapter. The coefficient indicates the impact of the variable on the roughness ranking. Simply put, the more positive a coefficient is, the more it is associated with rougher-than-average roughness rankings; the more negative a coefficient is, the more it is associated with smoother-than-average roughness rankings.

The model indicated that some individual attributes were associated with a higher likelihood of perceiving rough roads. The most important finding of this study may be that roughness, measured through IRI, was the most statistically significant factor associated with road roughness. In short, a rough test segment was highly correlated with a high roughness ranking, as one would expect. In addition, the age of the roadway surface and the subsequent presence of patchwork were associated with higher roughness ratings. The presence of expansion joints or bridge abutments was also significant, but the variable coefficient associated with patchwork was higher than that of joints/abutments, indicating that patchwork may be more important to drivers than joints/abutments.[20]

The analysis of test vehicle and roadway characteristics produced some interesting results. Sport utility vehicles and minivans were associated with lower roughness rankings. Also, increased noise inside the vehicle was associated with higher roughness rankings. While direct cause-effect inferences cannot be made, increased

[20] A helpful discussion of pavement defects is available in the *WSDOT Pavement Surface Condition Rating Manual* (Northwest Pavement Management Systems Users Group and Kay, 1992).

noise could influence how drivers perceive road roughness. In other words, because of more noise, drivers may perceive the road to be rough when it may not be. The model also indicated that speed was negatively associated with road roughness.

While they bear little use to transportation engineers, individual factors also influenced perceptions of road roughness. In general, women and older individuals (age 55 or over) were associated with lower (smoother than average) roughness rankings, while higher per capita household income was associated with higher (rougher than average) roughness rankings.

Frequent users of I-405 were associated with higher roughness rankings. However, IRI readings on I-405 were some of the lowest of all of the test segments. Equally inexplicable is the fact that frequent users of SR 520 were associated with lower roughness rankings. Table 10 shows that SR 520 was smoother on average (as measured by IRI), so it is possible that frequent users may have had a more positive predisposition to local highways. While it is difficult to explain why these variables were significant, it could be that they actually captured some mean unobserved effect shared by the individuals who "always" or "often" used these facilities.

Note that some variables were statistically significantly correlated. The age of the surface was highly correlated with IRI (Pearson Correlation of 0.745), as well as the presence of patchwork (Pearson Correlation of 0.431) and noise (Pearson Correlation of 0.284), as one might expect. The implications of correlated variables are discussed as part of the study's limitations at the end of this report.

BINARY LOGIT MODEL OF SEGMENT ROUGHNESS ACCEPTABILITY

The second model (Table 23) was more simplified because there was less variance to model in a dependent variable with a simple yes or no response. As a result, the model was expected to have fewer parameters to account for the lower variance. Note that because the dependent variable was "acceptability" (with 0 being "no" and 1 being "yes"), the signs in this model were the opposite of those found in Table 22.

Table 23. Binary Logit Model Segment Acceptability

Independent Variable		Estimated Coefficient	t-Statistic
Constant		4.496	21.890
IRI Measurement (in/mi) of roadway segment		-0.1563	-10.337
Age of Roadway Segment Surface (years)		-0.3152	-5.317
Number of Observations	2180		
Initial Log-Likelihood	-916.4279		
Log-Likelihood at Convergence	-703.4357		
Chi-Squared Statistic	425.9843		
Percentage Correctly Predicted	81.456		

Notes: Alternatives for the dependent variable were no (0) or yes (1). A positive coefficient means that the respondent was more likely to give an "acceptable" evaluation.

Most notable in this binary model may be what is excluded from the model. This model specification indicates that many factors that were significant in the roughness ranking were no longer significant in the acceptability decision. Vehicle type, noise, the presence of expansion joints/abutments, the presence of patch work, and level of service are just a few of the variables that were not statistically significant enough to be included in this model. In general, only the most important pavement-specific variables remained in the acceptability model.

As expected, IRI remained the most significant factor, and any segment with a high IRI was more likely to be considered unacceptable.[21] Test segments with older surfaces were also associated with lower acceptability evaluations, as might be expected.

ANALYSES OF VARIANCE

In this study, ANOVA techniques were used to validate the data collection process (i.e., testing whether factors that should be insignificant were indeed insignificant) and to support previously discussed results from the ordered probability models.

Analyses of Driving Direction and Road Facility on Mean IRI

Using the IRI data from the different directions back in Table 11, a simple ANOVA was used like a two-sample t-test to determine whether the measured IRI in the clockwise direction was statistically different from the roughness in the counterclockwise direction. The F-statistic of 1.157 indicates that there was not a statistically significant difference between directions at the 95 percent significance level, as shown in Table 24. Similarly, the IRI data on each of the four facilities (I-5, I-90, I-405 and SR 520) in Table 10 were tested for statistically significant differences. An F-statistic of 2.604 indicates that the differences in IRI on the four facilities were not statistically significant at the 95 percent significance level, as shown in Table 25.

---

[21] In the two models presented in this chapter, PSC was also tested in lieu of IRI as a better objective measure, but the presence of IRI in the model resulted in a higher t-statistic and better model fit than PSC.

Table 24. ANOVA for IRI Between Different Direction Groups

	Sum of Squares	df	Mean Square	F- Stat	Sig.
Between Groups	0.918	1	0.918	1.157	0.289
Within Groups	30.153	38	0.794		
Total	31.071	39			

Table 25. ANOVA for IRI Between Different Facility Groups

	Sum of Squares	df	Mean Square	F-Stat	Sig.
Between Groups	5.540	3	1.847	2.604	0.067
Within Groups	25.532	36	0.709		
Total	31.071	39			

Analyses of Start Location and Vehicle Type on Roughness Ranking, Acceptability, and Speed

In Table 26 and Table 27, the mean values for speed, roughness ranking, and segment acceptability are shown for different test vehicle types and different start locations, respectively.[22] These differences were tested for statistical significance.

Table 26. Mean Roughness and Acceptability Rankings and Speed Measurements by Test Vehicle

Vehicle	Speed (mph)	Roughness Ranking	Accept (N = 0, Y = 1)
Sedan (N = 711)	53.47 (12.76)	2.61 (1.17)	0.82 (0.39)
Sport Utility (N = 623)	56.35 (9.19)	2.52 (0.98)	0.83 (0.37)
Pick-Up (N = 459)	51.01 (12.83)	2.68 (1.09)	0.85 (0.36)
Minivan (N = 387)	57.39 (10.77)	2.59 (1.13)	0.82 (0.38)
Total (N = 2180)	54.47 (11.73)	2.60 (1.09)	0.83 (0.38)

Notes: Table contains mean values with standard deviations in parentheses.
Sample sizes may vary among some variables due to missing data.

[22] Speed is used to characterize changes in driver behavior. As will be shown later, speed should probably not be used as a continuous proxy for roughness ranking.

Table 27. Mean Roughness and Acceptability Rankings and Speed Measurements by Start Location

Start Location	Speed (mph)	Roughness Ranking	Accept (N= 0, Y = 1)
1 (N = 471)	54.22 (11.12)	2.38 (1.09)	0.89 (0.31)
2 (N = 612)	55.35 (12.24)	2.60 (1.11)	0.83 (0.38)
3 (N = 547)	51.81 (13.17)	2.56 (1.06)	0.85 (0.36)
4 (N = 550)	56.34 (9.47)	2.82 (1.06)	0.76 (0.43)
Total (N = 2180)	54.47 (11.73)	2.60 (1.09)	0.83 (0.38)

Notes: Table contains mean values with standard deviations in parentheses.
Sample sizes may vary among some variables due to missing data.

In Table 28, the F-statistics for the start location are insignificant. Start location should not affect driver behavior or the outcome of the roughness rankings, and these statistics support this assumption. At the same time, the F-statistics are significant for the test vehicle type, indicating that the hypothesis that “vehicle type does not affect the outcome of the roughness rankings” cannot be rejected.

Table 28. ANOVA of Start Location and Test Vehicle Type on Mean Speed

Source	DF	Sum of Squares	Mean Square	F-Stat	Pr > F
Start	3	24.9136	8.3045	0.28	0.8368
TestVeh	3	372.1116	124.0372	4.24	0.0116
Start*TestVeh	9	130.0386	14.4487	0.49	0.8691

Note: Dependent variable is the speed observed during roughness rankings.

The Duncan Grouping of speed and speed deviation by start location in Table 29 also supports the hypothesis that any observed differences in speed (or its variance) were not attributable to the start location.

Table 29. Duncan Multiple Comparison Test of Speed and Speed Deviation by Start Location

Location	Speed		Speed Deviation	
	Mean	Grouping	Mean	Grouping
U-Village Clockwise (N = 12)	54.223	A	8.895	A
U-Village CounterClockwise (N = 14)	55.207	A	9.400	A
Factoria Clockwise (N = 12)	51.704	A	11.281	A
Factoria CounterClockwise (N = 14)	56.349	A	8.359	A

Note: Means with the same letter grouping are not significantly different.

Table 30 supports the finding from the logit model that the drivers of the minivan and the pickup exhibited different travel behavior. Drivers of the minivan drove statistically significantly faster than those in the sedan and those in the pickup. At the same time, the pickup was driven statistically significantly slower than the SUV and minivan. (The mean speed exhibited in the SUV was not significantly different from the speeds exhibited in minivan or the sedan; the mean speed exhibited in the sedan was not significantly different from speeds exhibited in the SUV or the pickup.)

Table 30. Duncan Grouping of Speed and Speed Deviation by Test Vehicle Type

Test Vehicle	Speed		Speed Deviation	
	Mean	Grouping	Mean	Grouping
Minivan (N = 9)	58.177	A	7.843	A
SUV (N = 14)	56.753	A & B	6.887	A
Sedan (N = 18)	53.480	B & C	10.516	A & B
Pickup (N = 11)	50.193	C	12.222	B

Note: Means with the same letter grouping are not significantly different.

SUMMARY OF RESULTS

Drivers were placed in real world driving scenarios and asked to reveal their opinions about pavement roughness. This chapter identified some of the factors that were associated with changes in perceived roughness. With a combination of individual data, pavement data, and revealed roughness rankings, ordinal and binary logit models were estimated. The models probed the factors associated with differences in driver roughness rankings and roadway roughness acceptability. Both models showed vast improvements to the log-likelihood function and seemed to yield favorable goodness-of-fit statistics. The findings indicated that, while measured IRI levels provided the strongest indication of rankings and acceptability (as one would expect), other factors associated with the roadway, the test vehicle, and the individual were significant. In the next chapter, these results are compared with previous research to identify supporting and contrary findings.

CHAPTER 6. COMPARISON OF RESULTS

This chapter provides a brief comparison of the results from this research with past research. Individual, vehicle, and roadway characteristics significant to driver's perception of roughness are also reviewed.

In general, this research is difficult to compare to past research. First of all, few studies actually tested for significant factors; as mentioned in the literature review, many studies focused primarily on developing predictive equations for present serviceability rating (PSR) on the basis of roughness measurements. Second, of those studies that did test for significant associations with rankings, few studies included as many variables as were included in this study. In particular, most omitted socio-demographic variables because they were deemed to be useless in terms of applicability, as noted in the literature review of Garg et al. (1988). While it seems plausible that differences in perceived roughness among individuals may be influenced by their socio-demographic factors, many studies ignored their impact. The last reason for the difficulty in comparing results is that many results are in terms of PSR values, which are not directly translatable into IRI roughness rankings. Simply put, even though roughness rankings are similar to serviceability ratings, they are not identical and results cannot be directly compared. For example, when past researchers identified levels of acceptability in terms of PSR values, the only way that the research could be compared was possibly by visual inspection.

For the most part, results from this research coincide well with the results from past research, shown in Table 31. All studies, going back to the original AASHO Road Test, have found that roughness (however it is measured) is among the most statistically significant factors associated with roughness rankings (or ratings). Other factors that have been found to be significant and in agreement with past research include vehicle type and the presence of maintenance (i.e., "patch work"). While past studies had found age and gender to be insignificant, this research indicated that they correlated with roughness rankings.

Table 31. Comparison of Findings with Past Research

Significant	Not Significant
• Surface roughness (+: N, FL) • Vehicle type (+: N) • Pavement maintenance (+: N) • Rater age (-: N, NJ) • Rater gender (-: N, NJ) • Vehicle speed (-: N)	• Pavement type (-: N)

Key: (+) agrees with past finding, (-) conflicts with past finding
Past Findings: NJ = Nick & Janoff (1983); J = Janoff et al. (1985);
N = Nair et al. (1985); FL = Fernando & Lee (1999).

One of the most interesting findings is that vehicle speed is significant, which conflicts with the findings of past research. This finding is significant because other limited research has seemed to suggest that drivers slow down as road roughness increases (e.g., Paterson and Watanatada, 1985). Note that many past studies kept the vehicle speed constant and did not treat it as a variable.

Another interesting finding is that pavement type is not significant. This conflicts with past research and, admittedly, seems counterintuitive. Because rigid and flexible pavements are so different structurally, one might expect associated differences in roughness rankings. It is possible that the limited amount of flexible pavement in this study may have had an impact.

Another useful comparison is among acceptability threshold results. Unfortunately, as mentioned earlier, many of these results are in terms of PSR and are difficult to compare. Still, it is interesting to point out that while Nair et al. (1985) concluded that “up to 88 percent of the variation in PSR can be explained by the roughness variables” (p. 81), this study concluded that about 83 percent of the variability can be explained by a simple linear representation of roughness rankings (see Figure 13). While Nair et al. concluded that their study “further attests to the serviceability-performance concept in general and to the validity of using road profile measurements to predict PSRs and to obtain indices of serviceability in particular” (p. 163), this study indicates that IRI, which is an easier measure to quantify and obtain, is also a good predictor of roughness rankings.

CHAPTER 7. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

This study identified salient factors associated with drivers' perceptions of road roughness on urban highways. From the data collection process to the analysis process, the focus moved beyond the present-serviceability concept. Even though indexed present serviceability ratings (PSRs) may be a good measure of an individual's perception of roughness rankings, this research brought more easily quantifiable measures to the forefront of road roughness assessment, particularly the International Roughness Index (IRI). IRI is a widely accepted, standardized, objective measure that was shown to be the most significant factor associated with changes in drivers' perceptions of road roughness, as well as drivers' acceptability of a roadway's condition.

IRI, PSC, and rutting depth are measures that are maintained and monitored in the state pavement management system (WSPMS) database. All are tools for highway engineers and preservation program decision-makers.

SUMMARY OF MAJOR FINDINGS

Patchwork and bridge abutments connecting bridge or elevated facilities are associated with rougher-than-average rankings by drivers. While some drivers may understand that these imperfections are necessary parts of the highway system, especially around Seattle with many elevated sections, this research suggests that these imperfections affect drivers' perceptions of road roughness. The models also suggested that the factors listed in Table 32 affect drivers' perceptions of roughness. While transportation engineers have little or no control over some of these factors (e.g. user socio-demographics), it is important to recognize the differences among users because these differences could account for some of the variability observed in roughness rankings.

Table 32. Summary of Significant Variables

Variables Associated with More Roughness	Variables Associated with Less Roughness
• Measured IRI • Observable "maintenance" • Presence of joints/abutments • Age of surface • In-vehicle noise • Vehicle speed • High income users • Male users • Frequent users of I-405	• Older individuals • Sport utility test vehicles • Minivan test vehicles • Female users • Frequent users of SR 520

This study also provided empirical data that can be used to support an IRI acceptability threshold. The IRI guideline recommended by the 1998 *National Strategic Plan* set forth by FHWA is 170 in/mi (2.7 m/km), and this study found that approximately 85 percent of all "acceptable" evaluations fell at or below this value.

LIMITATIONS

As with any research study, it is always beneficial to alert the reader to any limitations or hazards that may be inherent in the results. Whenever data are pooled together from multiple sources, caution must be taken. Furthermore, as with most research that deals with the human element, this study has its limitations.

First of all, potential statistical biases are inherent in some of the analytical methods. With multiple observations from single respondents and results based on multiple segments on the same roadway facility with similar attributes, there is a potential for multiple cases of positive serial correlation within the data, which could invalidate some of the logit model results. Serial correlation, if present when independent variables are related, can either be minor or can underestimate the model's standard errors and yield misleading levels of significance. In other words, factors could appear to be more significant than they actually are.

Given this potential pitfall, care was taken to make conclusions about variables with a significance level beyond what would normally be accepted. In other words, to help ensure that variables were significant, they were held to a higher level of significance than normal.

Other limitations include the fact that results were based on pavements with a limited range of IRI values. A wider range of "poor" pavements may have provided more useful results. Unfortunately, the project was limited to the accessibility of facilities around the Seattle area and to individuals who were willing to participate.

Another limitation was the fact that drivers had to make a single, discrete decision about a continuous test segment. Because test segments were not completely homogenous throughout the section, drivers had to use their best judgment when evaluating test segments. In an ideal testing scenario, it would be better if each test segment were completely homogenous, but this is not realistic.

One inherent problem that was observed in the data collection was that some individuals may have made unfair comparisons between local city streets and state highway facilities. While it is understandable that the general public holds local city streets and state-managed highways to similar standards of service, most of the general public may not realize that these separate facilities are managed and maintained by different jurisdictions and at different standards. As a result, individuals may have made subconscious comparisons between highways and city streets, which may have affected their evaluations of the highway test sections.

In addition to preconceived biases, there was also an inherent self-selection among participants, which may have affected the results. The people who participated in the in-vehicle surveys selected themselves by contacting the researchers and scheduling an appointment. Participants had to return the preliminary survey with their name and contact information before they could even be considered to participate in the in-vehicle data collection. As noted earlier, another limitation of the driver sample is that they were selected from the traffic stream close to the University of Washington campus and may not represent the typical users of the Washington State highway system.

REFERENCES

Arterburn, S. and B. Suprenant (1990). P*ublic Perception of Pavement Rideabilit*y," Colorado Department of Highways. Report CDOH-UCB-R-90-10. Denver, Colorado. July.

Boeger, R. and R.J. Crowe (2002). "It's the Ride that Counts." *Public Road,* 65 (4). January/February. pp. 17-19. Available on-line at: *http://www.tfhrc.gov/pubrds/janfeb02/maricopa.htm*.

Cambridge Systematics (1998). *Department of Transportation Highways and Rail Programs Performance Audit*. Prepared for the Washington State Joint Legislative Audit and Review Committee (JLARC). Report 98-2. Available on-line at: *http://jlarc.leg.wa.gov/Reports/98-2HwyRails.PDF*.

Carey, W.N., Jr. and P.E. Irick (1960). The Pavement-Serviceability Concept. *Highway Research Board Bulletin 250*. National Research Council. Washington, D.C., pp 40-58.

Coopers and Lybrand, L.L.P. (1995). National Quality Initiative (NQI) National Highway User Survey (NHUS). Federal Highway Administration, Washington, D.C. Available on-line at: http://ntl.bts.gov/data/nqinhus.pdf.

Elway Research Inc. (1997). Transportation Agencies Performance Audit: Citizen Input Survey. Prepared for the Washington State Joint Legislative Audit and Review Committees (JLARC). September.

Federal Highway Administration (2001). 1999 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance, Report to Congress. Federal Highway Administration. Report FHWA/PL-08/017. Washington DC. May. Available on-line at: http://www.fhwa.dot.gov///////policy/1999cpr/index.htm.

Fernando, E. and N.Y. Lee (1999). Technical Memorandum. "Documentation of Ride Surveys Conducted in Project 4901." Texas Transportation Institute. Texas A&M University System. College Station, TX. December.

Garg, A, A. Horowitz, and F. Ross (1988). "Establishing Relationships Between Pavement Roughness and Perceptions of Acceptability," *Transportation Research Record 1196*, Transportation Research Board, National Research Council, Washington, D.C.

Gillespie, T.D. (1992). "Everything You Always Wanted to Know about the IRI, But Were Afraid to Ask!" Presented at the Road Profile Users Group Meeting, Lincoln, Nebraska, September 22-24. Available on-line at: *http://www.umtri.umich.edu/erd/roughness/every.pdf*.

Gillespie, T.D., M.W. Sayers, and L. Segel (1980). Calibration and Correlation of Response-Type Road Roughness Measuring Systems. *National Cooperative Highway Research Program Report 228*, Transportation Research Board, National Research Council, Washington D.C. December.

Greene, W.H. (1993). Econometric Analysis. Fourth Edition. Macmillan Press. New York.

Hicks, C. (1993). Fundamental Concepts in the Design of Experiments. Fourth Edition. Oxford University Press.

Holbrook, L.F. (1969). "Prediction of Subjective Response to Road Roughness by use of the Rapid Travel Profilometer." *Highway Research Record 69*. National Research Council. Washington, D.C., pp 212-226.

Hudson, W.R. (1981). "Road Roughness: Its Elements of Measurement," *Transportation Research Record 836*, Transportation Research Board, National Research Council, Washington, D.C.

Hutchinson, B.G. (1963). "Principles of Subjective Rating Scale Construction," *Highway Research Record 46*, National Research Council. Washington, D.C., pp. 60-70.

Janoff, M.S. and J.B. Nick (1983). "Effects of Vehicle and Driver Characteristics on the Subjective Evaluation of Road Roughness" in Measuring Road Roughness and Its Effects on User Cost and Comfort. *ASTM Special Technical Publication (STP) 884*, T.D. Gillespie and M. Sayer, Eds. American Society for Testing and Materials, Philadelphia. pp. 111 - 126.

Janoff, M.S., J.B. Nick, P.S. Davit, and G.F. Hayhoe (1985). Pavement Roughness and Rideability. *NCHRP Report 275*. National Cooperative Highway Research Program. Transportation Research Board, National Research Council, Washington, D.C. September.

Karan, M.A., D.H. Kobi, C.B. Bauman, R.C. Hass (1979). "Subjective and Mechanical Estimations of Pavement Serviceability for Rural-Urban Road Networks," *Transportation Research Record 715*, Transportation Research Board, National Research Council, Washington, D.C. pp. 31-36.

Kennedy, P. (1992). A Guide to Econometrics. Third Edition. MIT Press, Cambridge, MA.

Keever, D.B., K.E. Weiss, and R.C. Quarles (2001). *Moving Ahead: The American Public Speaks on Roadways and Transportation in Communities*. Federal Highway Administration (FHWA) Report: FHWA-OP-01-017. Washington, D.C.

40 pp. February. Available on-line at: http://www.fhwa.dot.gov/reports/movingahead.pdf.

Khattak, A., J. Schofer, and F. Koppelman (1993). "Commuters' Enroute Diversion and Return Decisions: Analysis and Implications for Advanced Traveler Information Systems." *Transportation Research 27A(2)*, pp. 101-111.

Kuemmel, D.A., R.K. Robinson, R.J. Griffin, R.C. Sonntag, and J.K. Giese (2001a). Public Perceptions of the Midwest's Pavements –Final Phase III Report – Iowa. Report CHTE 2001-02. WisDOT Highway Research Study #94-07. January. Available on-line at: *http://www.trc.marquette.edu/public_perception/*.

Kuemmel, D.A., R.K. Robinson, R.J. Griffin, R.C. Sonntag, and J.K. Giese (2001b). Public Perceptions of the Midwest's Pavements –Executive Summary – Minnesota. Report CHTE 2001-02. WisDOT Highway Research Study #94-07. February. Available on-line at: *http://www.trc.marquette.edu/public_perception/*.

Kuemmel, D.A., R.K. Robinson, R.J. Griffin, R.C. Sonntag, and J.K. Giese (2001c). Public Perceptions of the Midwest's Pavements –Executive Summary – Wisconsin. Report WI/SPR-01-01. WisDOT Highway Research Study #94-07. January. Available on-line at: *http://www.trc.marquette.edu/public_perception/*.

Moore, R.K., G.N. Clark, and G.N. Plumb (1987). "Present Serviceability–Roughness Correlations Using Rating Panel Data." *Transportation Research Record 1117*, Transportation Research Board, National Research Council, Washington, D.C. pp. 152-158.

Nair, S.K., W.R. Hudson, and C.E. Lee (1985). *Realistic Pavement Serviceability Equations Using the 690D Surface Dynamics Profilometer*. Research Report 354-1F, Center for Transportation Research, University of Texas, Austin. Austin, Texas. August. [Some results from this report are also available in: Nair, S.K. and W.R. Hudson (1986). "Serviceability Prediction From User-Based Evaluations of pavement Ride Quality," *Transportation Research Record 1084*, Transportation Research Board, National Research Council, Washington, D.C. pp. 66-72.]

Nakamura, V.F., and H.L. Michael (1963). "Serviceability Ratings of Highway Pavements." *Highway Research Record 40*. Highway Research Board. Washington, D.C., pp. 21-36.

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman (1990). Applied Linear Models. Fourth Edition. Irwin Press, Chicago.

Ng, L., W. Barfield, and F. Mannering (1995). "A Survey-Based Methodology To Determine Information Requirements For Advanced Traveler Information Systems." *Transportation Research 3C(2)*, pp. 113-127.

Nick, J.B. and M.S. Janoff (1983). "Evaluation of Panel Rating Methods for Assessing Pavement Ride Quality." *Transportation Research Record 946*, Transportation Research Board, National Research Council, Washington, D.C. pp. 5-13.

Northwest Pavement Management Systems Users Group and R.K. Kay (1992). Pavement Service Condition Rating Manual. Washington State Department of Transportation. March. Available on-line at: http://www.wsdot.wa.gov/TA/T2Center/Mgt.Systems/PavementTechnology/PSCRMan.pdf.

Paterson, W.D.O. and T. Watanatada (1985). "Relationships Between Vehicle Speed, Ride Quality, and Road Roughness" in Measuring Road Roughness and Its Effects on User Cost and Comfort. *ASTM Special Technical Publication (STP) 884*, T.D. Gillespie and M. Sayer, Eds. American Society for Testing and Materials, Philadelphia. pp. 89-110.

Sayers, M.W., T.D. Gillespie, and C.A.V. Queiroz (1986). "The International Road Roughness Experiment: A Basis for Establishing a Standard Scale for Road Roughness Measurements." *Transportation Research Record 1084*, Transportation Research Board, National Research Council, Washington, D.C. pp. 76-85.

Scrivner, F.H., and W.R. Hudson (1964). "A Modification of the AASHO Road Test Serviceability Index Formula." *Highway Research Record 46*. Highway Research Board. Washington, D.C. pp. 71-87.

Sivaneswaran, N., L.M. Pierce, D. Crimmins, and J.R. Livingston (2000). *WSPMS, Version 2000.1.0*. Washington State Department of Transportation, FOSSC Materials Laboratory. Olympia, Washington.

Transportation Research Board (2000). *Highway Capacity Manual 2000*. National Research Council, Washington, DC.

United States Census Bureau (unpublished). American FactFinder. Data available on-line at: http://factfinder.census.gov/.

Ward, D.R., K.J. Kercher, and S. Gulen (1993). *Correlation of IRI to Public Perception of Pavement Roughness: Final Report*. Indiana Department of Transportation, Division of Research. West Lafayette, Indiana. May.

Washington State Department of Transportation (2001). "Pavement Testing." Pavement Distress Identification Van picture courtesy of WSDOT Material Laboratory website. Available on-line at: *http://www.wsdot.wa.gov/fossc/mats/pavement/distress.htm*.

APPENDIX A. SAMPLE PRELIMINARY SURVEY

University of Washington

Road Quality and Highway Roughness Survey

The goal of this survey is to obtain your opinion of the quality and driving comfort of various highways in Washington State. Your views, experiences and insights will be greatly appreciated. It is hoped that this survey results will help the Washington State Department of Transportation prioritize future highway funding, so your participation and input will make a difference.

Please give this survey to the person in your household who most often drives. Please ask him/her to fill out the survey and return it by mail by ***December 15, 2000***. No postage is necessary. We appreciate your response. This survey is anonymous and your answers will not be associated with your name. If you have any questions, please feel free to call (206) 685-6779.

A. General Information

1. Do you regularly commute (for work) on Seattle-area roads? ☐ Yes ☐ No

2. How many people are usually in your vehicle when you travel? (check one only)

   ☐ 1 ☐ 2 ☐ 3 ☐ 4 or more

3. How congested do you think Seattle-area highways are?

   ☐ Heavy Congestion
   ☐ Moderate Congestion
   ☐ Mild Congestion
   ☐ Little Congestion
   ☐ No Congestion

4. What is your opinion of the roughness of Washington State Roads?

   (check **one**) ☐ Very Rough ☐ Rough ☐ Average ☐ Smooth ☐ Very Smooth

5. How would you rate the noise level in the vehicle you drive most often?

   (check **one**) ☐ Very Noisy ☐ Noisy ☐ Average ☐ Quiet ☐ Very Quiet

6. How would you rate the ride quality in the vehicle you drive most often?

   (check **one**) ☐ Very Rough ☐ Rough ☐ Average ☐ Smooth ☐ Very Smooth

7. How often do you use the following highways in a normal week? (Check **one** box per line.)

	Always use	Often use	Sometimes use	Seldom use	Never use
Inte					
Interstate 90	☐	☐	☐	☐	☐
Interstate 405	☐	☐	☐	☐	☐
State Route 520	☐	☐	☐	☐	☐

B. Your Opinions of Driving on Washington State Roads

In the vehicle you drive most often in the Seattle area, please indicate how smooth or rough you believe the following highway segments are:

						Very Rough
8. ***Interstate 5 in the area around:***						
a. SeaTac Airport	☐	☐	☐	☐	☐	☐
b. Downtown Seattle	☐	☐	☐	☐	☐	☐
c. N						☐
9. ***Interstate 90 in the area around***						
a. Mercer Island	☐	☐	☐	☐	☐	☐
b. B						☐
c. Is						☐
10. ***Interstate 405 in the area around*:**						
a. Southcenter	☐	☐	☐	☐	☐	☐
					☐	☐
c. B						☐
11. ***State Route 520 in the area around*:**						
a. M					☐	☐
b. B						☐

C. Yourself

12. Are you: ☐ Female ☐ Male

13. Are you: ☐ Married ☐ Single ☐ Other

14. What is your age?

- ☐ Under 21
- ☐ 21-25
- ☐ 26-30
- ☐ 31-35
- ☐ 36-40
- ☐ 41-45
- ☐ 46-50
- ☐ 51-55
- ☐ 56-60
- ☐ 61-65
- ☐ 66-70
- ☐ Over 70

15. What is your approximate annual household income?

- ☐ no income
- ☐ under $15,000
- ☐ $15,000-24,999
- ☐ $25,000-34,999
- ☐ $35,000-44,999
- ☐ $45,000-54,999
- ☐ $55,000-64,999
- ☐ $65,000-74,999
- ☐ $75,000-84,999
- ☐ $85,000-99,999
- ☐ $100,000 – 150,000
- ☐ over 150,000

16. What is your highest level of education?

- ☐ some high school
- ☐ high school diploma
- ☐ technical college degree (A.A.)
- ☐ college degree (Bachelors degree)
- ☐ post-graduate degree

17. In the past 5 years, how many traffic accidents have you been involved in? ___________

18. How often do you wear seatbelts while driving? (Check **one** box only)

☐ all of the time ☐ most of the time ☐ some of the time ☐ rarely ☐ never

19. What are the year, make, and model of the vehicle you drive most often on Seattle-area highways?

Year ____________ Make _______________ Model _______________

20. How many licensed and operable vehicles do you have in your household? ______________

21. Including yourself, how many people live in your household?_________

22. How many children under age 6 live in your household? _______________

23. How many children 6 to 16 live in your household? _______________

24. How many people living in your household work outside the home?_____

25. Would you be willing to participate in a driving experiment on highway roughness?

☐ Yes ☐ No

26. If you answered yes to question #25, please include your name, address and telephone number below so that we may contact you for further information and assistance.

____

____

____

THANK YOU FOR YOUR PARTICIPATION!

APPENDIX B. IN-VEHICLE PARTICIPATION LETTER

University of Washington

Road Quality and Highway Roughness Study

June 1, 2001

Dear <Participant Name>:

In December, you completed the Road Quality and Highway Roughness survey for research being done at the University of Washington. In completing the survey, you expressed a willingness to participate in a driving experiment.

This summer, the follow-up driving experiment will be conducted, and we need your help. The driving study is a simple experiment that asks you about your opinion of the quality around the Seattle-area freeways while you are driving. The experiment should take less than two hours to complete, and **you will receive $50 as compensation for your time**. If you participate in this experiment, an insured vehicle will be provided but you must have a valid Washington State driver's license.

If you would like to learn more about being part of this unique research opportunity, please call the University of Washington Department of Civil & Environmental Engineering at (206) 934-4899. Be sure to leave your name, phone number, and the best time to reach you. After you call, you can be scheduled to participate in this experiment at a time that is convenient for you – but space is limited.

Your participation in this study is very important to the State of Washington, as we hope it will result in improved road conditions that better suit your needs and better spend your highway tax dollars.

Thanks in advance for your help.

Sincerely,

Kevan Shafizadeh, Research Engineer
Dept. of Civil & Environmental Engineering
Box 352700, University of Washington
Seattle, WA 98195-2700

APPENDIX C. SUMMARY STATISTICS FROM SURVEYS

Table C-1. Summary Statistics of Survey Respondents

Variable	Preliminary	In-Vehicle
Gender	**Frequency (%)**	**Frequency (%)**
Female	207 (39.8%)	23 (41.1%)
Male	313 (60.2%)	33 (58.9%)
Marital Status	**Frequency (%)**	**Frequency (%)**
Married	322 (63.4%)	32(57.1%)
Single	163 (32.1%)	18 (32.1%)
Other	23 (4.5%)	6 (10.7%)
Age	**Frequency (%)**	**Frequency (%)**
Less Than 21	2 (0.4%)	0 (0.0%)
21 - 25	29 (5.6%)	6 (10.7%)
26 - 30	47 (9.1%)	6 (10.7%)
31 - 35	69 (13.3%)	5 (8.9%)
36 - 40	57 (11.0%)	5 (8.9%)
41 - 45	66 (12.7%)	5 (8.9%)
46 - 50	73 (14.1%)	8 (14.3%)
51 - 55	73 (14.1%)	7 (12.5%)
56 - 60	39 (7.5%)	8 (14.3%)
61 - 65	24 (4.6%)	3 (5.4%)
66 - 70	13 (2.5%)	1 (1.8%)
Over 70	26 (5.0%)	2 (3.6%)
Income	**Frequency (%)**	**Frequency (%)**
No Income	2 (0.4%)	1 (2.0%)
Less than $15,000	4 (0.9%)	2 (3.9%)
$15,000 - $24,999	14 (3.0%)	2 (3.9%)
$25,000 - $34,999	33 (7.1%)	0 (0.0%)
$35,000 - $44,999	31 (6.7%)	8 (15.7%)
$45,000 - $54,999	40 (8.6%)	7 (13.7%)
$55,000 - $64,999	48 (10.3%)	3 (5.9%)
$65,000 - $74,999	42 (9.1%)	6 (11.8%)
$75,000 - $84,999	45 (9.7%)	4 (7.8%)
$85,000 - $99,999	55 (11.9%)	4 (7.8%)
$100,000 - $150,000	89 (19.2%)	13 (25.5%)
Over $150,000	61 (13.1%)	1 (2.0%)
Education	**Frequency (%)**	**Frequency (%)**
Some High School	3 (0.6%)	0 (0.0%)
High School Diploma	65 (12.6%)	5 (9.1%)
Associates Degree	44 (8.5%)	7 (12.7%)
College Degree	229 (44.5%)	26 (47.3%)
Post-Graduate Degree	174 (33.8%)	17 (30.9%)

Variable	Preliminary	In-Vehicle
Regular Commute	**Frequency (%)**	**Frequency (%)**
No	84 (16.2%)	7 (12.5%)
Yes	437 (84.0%)	49 (87.5%)
Vehicles in Household	**Frequency (%)**	**Frequency (%)**
0	0 (0.0%)	1 (1.8%)
1	137 (26.2%)	11 (19.6%)
2	238 (45.6%)	29 (51.8%)
3	90 (17.2%)	8 (14.3%)
4	44 (8.4%)	4 (7.1%)
5	10 (1.9%)	2 (3.6%)
6	1 (0.2%)	0 (0.0%)
7	1 (0.2%)	0 (0.0%)
8+	1 (0.2%)	1 (1.8%)
Infants in Household	**Frequency (%)**	**Frequency (%)**
0	465 (89.3%)	51 (91.1%)
1	41 (7.9%)	3 (5.4%)
2	13 (2.5%)	1 (1.8%)
3	2 (0.4%)	1 (1.8%
Children in Household	**Frequency (%)**	**Frequency (%)**
0	403 (77.4%)	47 (83.9%)
1	59 (11.3%)	5 (8.9%)
2	42 (8.1%)	3 (5.4%)
3	14 (2.7%)	1 (1.8%)
4+	2 (0.4%)	0 (0.0%)
Workers in Household	**Frequency (%)**	**Frequency (%)**
0	53 (10.2%)	2 (3.6%)
1	188 (36.2%)	19 (34.5%)
2	217 (41.8%)	27 (49.1%)
3	45 (8.7%)	6 (10.9%)
4	11 (2.1%)	0 (0.0%)
5	2 (0.4%)	0 (0.0%)
6+	3 (0.6%)	1 (1.8 %)
Household Size	**Frequency (%)**	**Frequency (%)**
0	2 (0.4%)	0 (0.0%)
1	93 (17.8%)	7 (12.5%)
2	214 (40.9%)	28 (50.0%)
3	87 (16.6%)	10 (17.9%)
4	82 (15.7%)	6 (10.7%)
5	28 (5.4%)	1 (1.8%)
6	12 (2.3%)	3 (5.4%)
7	2 (0.4%)	0 (0.0%)
8+	3 (0.6%)	1 (1.8%)

APPENDIX D. SUMMARY OF DRIVER SEGMENT PROFILES

This table that summarizes some of the roadway characteristics of the different test segments. The segments labeled C1 through C20 are the first through twentieth segments in the clockwise segments, respectively, and the test segments numbered CCW1 through CCW20 are the first through twentieth segments in the counterclockwise direction.

Sect	SR	Side	Lane	BSRMP	ESRMP	Lane	IRI	Terrain	LtShld	RtShld	Lanes	Width	Surf	Age	Bridge
C1	520	I	R	1.20	1.56	R	3.01	L	0	0	2	24	P	37	B
C2	520	I	R	1.63	2.19	R	1.46	L	0	0	2	24	A	4	B
C3	520	I	R	3.00	3.46	R	1.33	L	0	0	2	24	A	4	B
C4	520	I	R	4.26	5.17	R	1.24	R	0	0	2	24	A	4	N
C5	520	I	R	5.26	5.78	R	1.22	R	8	0	2	24	A	4	N
C6	405	D	R	14.38	14.14	R	0.98	R	10	0	3	36	A	7	N
C7	405	D	R	14.01	13.82	R	1.02	R	10	0	3	36	A	7	N
C8	405	D	R	13.68	13.55	R	1.17	R	10	0	3	36	A	7	N
C9	405	D	R	13.41	13.10	R	2.00	R	10	0	3	36	A	6	N
C10	405	D	R	12.65	12.33	R	1.95	R	10	10	3	36	P	27	N
C11	090	D	R	8.26	7.86	R	1.55	R	10	10	3	36	P	9	N
C12	090	D	R	7.18	6.62	R	1.60	R	10	6	3	36	P	9	N
C13	090	D	R	6.56	6.03	R	1.43	R	0	0	3	44	P	12	N
C14	090	D	R	5.58	4.54	R	1.15	R	0	0	3	44	P	16	B
C15	090	D	R	4.23	3.57	R	1.56	R	0	0	3	44	P	14	N
C16	090	D	R	3.25	3.08	R	1.89	R	10	10	4	48	P	8	N
C17	005	I	R	165.67	166.15	R	2.70	R	0	0	4	64	P	36	B
C18	005	I	R	166.20	166.91	R	2.59	R	6	12	4	48	P	36	N
C19	005	I	R	167.02	167.35	R	3.47	R	0	0	4	60	P	36	B
C20	520	I	R	0.33	0.56	R	2.27	L	0	0	2	26	P	40	B
CCW1	520	D	R	0.74	0.30	R	3.06	L	0	0	2	26	P	40	B
CCW2	005	D	R	166.70	166.37	R	1.68	R	0	0	4	60	P	37	B
CCW3	005	D	R	166.18	165.83	R	2.36	R	10	2	4	48	P	34	N
CCW4	090	I	R	2.97	3.07	R	1.62	R	8	10	4	48	P	12	N
CCW5	090	I	R	3.83	4.20	R	2.88	R	0	0	3	47	P	61	N
CCW6	090	I	R	4.46	5.53	R	1.44	R	0	0	3	56	P	8	B
CCW7	090	I	R	6.06	6.54	R	1.30	R	0	0	3	54	P	12	N
CCW8	090	I	R	6.67	7.24	R	1.74	R	6	10	3	36	P	9	N
CCW9	090	I	R	7.41	8.14	R	1.63	R	6	10	3	36	P	9	N
CCW10	405	I	R-1	12.05	12.39	R-1	3.76	R	6	10	3	36	P	29	N
CCW11	405	I	R-1	12.48	12.73	R-1	2.58	R	6	10	3	36	P	29	N
CCW12	405	I	R-1	13.01	13.30	R-1	4.28	R	0	10	3	36	P	29	N
CCW13	405	I	R	13.55	13.97	R	3.51	R	0	10	3	36	P	29	N
CCW14	520	D	R-2	5.68	5.30	R-2	0.89	R	8	0	2	24	A	4	N
CCW15	520	D	R-2	5.15	4.71	R-2	1.00	R	0	0	2	24	A	4	N
CCW16	520	D	R-2	4.61	4.31	R-2	1.18	R	0	0	2	24	A	4	N
CCW17	520	D	R	3.61	3.16	R	1.07	L	0	0	2	24	A	4	B
CCW18	520	D	R	3.01	2.56	R	1.06	L	0	0	2	24	A	4	B
CCW19	520	D	R	2.23	1.68	R	1.33	L	0	0	2	24	A	4	B
CCW20	520	D	R	1.53	1.41	R	3.28	L	0	0	2	24	P	37	B

APPENDIX E. DISCUSSION OF IN-VEHICLE SURVEY INSTRUMENTS

The in-vehicle measurement instrument has been somewhat contested in past research and deserves a brief discussion, in part, because it challenges some of the fundamental data collection that was used the performance-serviceability concept. This appendix reviews the evolution of the passenger panel-rating instrument and shows how it lends itself to the simplified individual driver-rating instrument.

The original AASHO rating scale is shown in Figure E-1, but it was criticized by researchers, such as Holbrook and Weaver, and was modified into Weaver/AASHO scale, which contained anchors on each end of the scale.

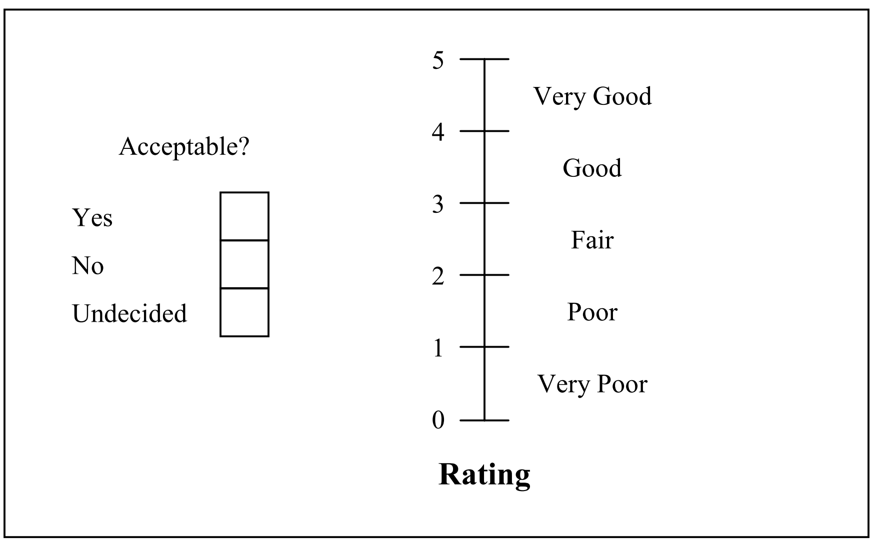


Figure E-1. The Original AASHO Passenger Panel Rating Scale

The modified AASHO/Weaver rating scale, shown in Figure E-2, never gained acceptance. Instead, researchers reverted back to the original AASHO rating scale. One possible explanation may have been that, by this point, the original instrument was already ingrained in practice. Or, it could be that researchers did not support its original criticism and wanted to compare results with the original AASHO Road Test.

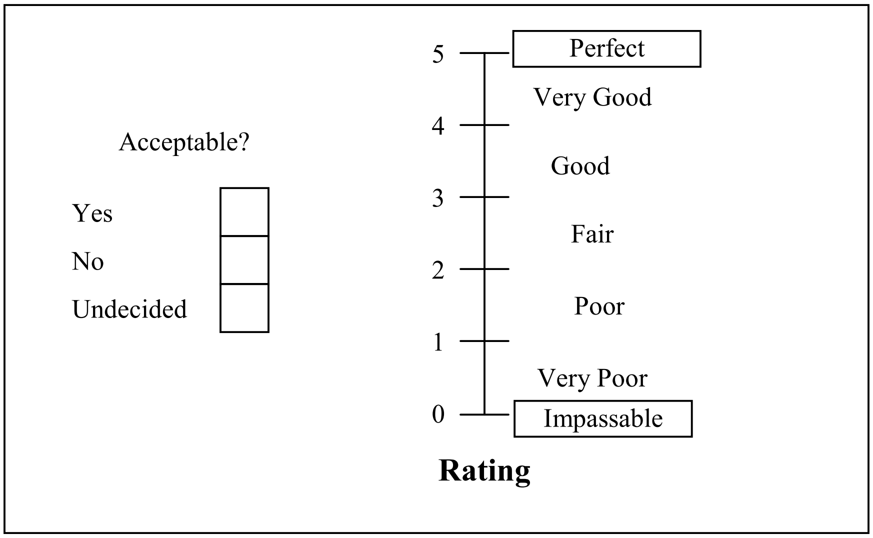


Figure E-2. The Modified AASHO/Weaver Passenger Panel Rating Scale

Using a scale similar to the one shown in Figure E-1, researchers converted numerical values to marks made by passenger evaluators by assuming that the scale was continuous (i.e., not discrete) that there was a linear relationship between ratings. Even though the roughness rankings are somewhat subjective, the conversion from panel markings to numerical values was done with unrealistic precision. While Fernando and Lee (1999) divided the scale into "tenths of a rating point," Nair et al. (1985) converted the marks on the rating panel to a numerical value precise to a "hundredths of a rating point":

> Ratings were scaled off the rating forms to the nearest 0.01 point by using a ruler placed alongside the vertical scale. These ratings, by definition, are the individual present serviceability ratings (IPSRs). The mean of these IPSRs for a section becomes the present serviceability rating (PSR) for that section. (p. 65, Nair et al, 1985).

Hutchinson (1963) and Holbrook (1969) studied the affect of different rating instruments. In 1963, Hutchinson reviewed the AASHO rating scale its potential subversion by various distortions and biases, such as the error of leniency, the halo effect, central tendency, and anchoring.[23] The Holbrook scale, shown in Figure E-3, is also a direct-type panel rating instrument and contains specific cues along the scale to guide the evaluator. The problem with this instrument is that the cues can be associated with unexpected connotations that could lead to biased evaluations.

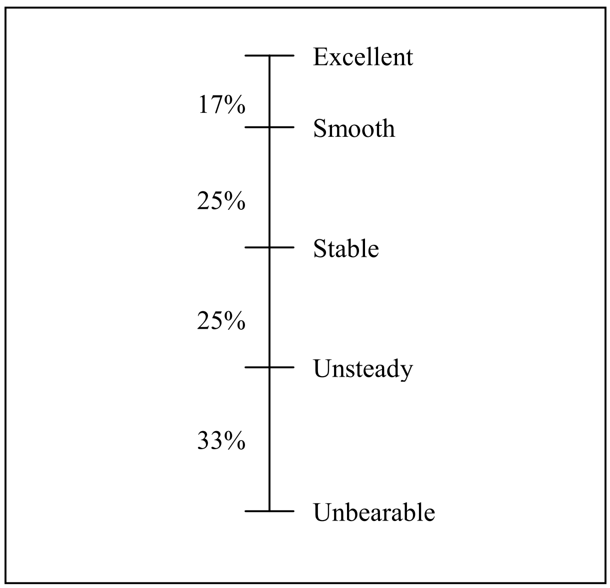


Figure E-3. The Holbrook Passenger Panel Rating Scale

The non-segmented rating scale, shown in Figure E-4, is a direct scale that eliminates the biases caused by cue words. The disadvantage of this scale is that respondents have difficulty using it.

[23] The error of leniency refers to the tendency of a rater to rate too high or too low for any reason. The halo effect refers to the tendency of a rater to force the rating of a particular attribute in the direction of the overall impression of the object rated. The error of central tendency refers to the tendency of a rater to hesitate in giving extreme ratings and to make ratings toward the average of the group.” (Hutchinson, 1964, pp. 63-64)

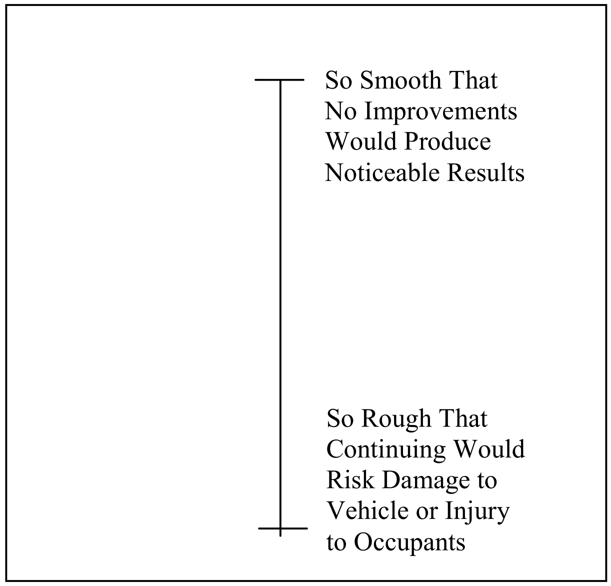


Figure E-4. The Non-Segmented Passenger Panel Rating Scale

The successive categories method is an indirect method that relies on the individual to make rational judgments. It is easy to complete verbally, especially while driving. Its primary drawback is that the values obtained from this method are difficult to compare to other scale values, such as the PSR values. This survey method is effective, however, because it captures major changes in driver perception and forces the driver to make rational decisions about discrete values. The same can be said with the acceptability response. The "undecided" choice was eliminated because it forced the respondent to make difficult decisions.

Each of the three rating instruments was evaluated and after statistical analyses were performed, it was concluded that there was no significant difference between the three scales (i.e., R^2 values ranged from 0.844 to 0.899) (Janoff et al, 1985). Other advantages and disadvantages are highlighted in Table E-1.

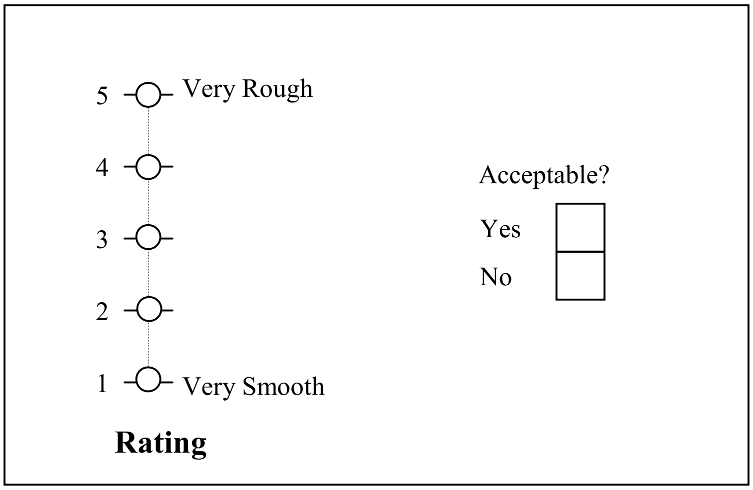


Figure E-5. Schematic of Individual Driver Rating Scale

Table E-1. Comparison of Scaling Methods

Method	Advantages	Disadvantages
Weaver/AASHO	None	Subject to measurement errors.
Holbrook's Scale	Accurate placement of cues along the scale should aid the observers in making direct interval judgments	Connotative problems associated with intermediate cue words could bias results.
Non-Segmented Scale	Eliminates any problems introduced using intermediate cue words	Many observers may find it difficult to make their ratings without the aid of cue words.

Source: Janoff et al. (1985).

Name of Participant: ___________________________________ Date: _______________
Start Location and Direction: ____________________________
Weather Conditions: ____________________ Vehicle Type: ______________________

Segment #	Wet or Dry Pavement?	Noise (dB)	Speed (mph)	LOS (A,B,C,D,F)	Roughness (1 = smooth, 5 = rough)	Acceptable? (Yes or No)

University of Washington Civil & Environmental Engineering Pavement Quality and Road Roughness Study

Figure E-6. In-Vehicle Driver Survey Worksheet (Completed by Researcher)

APPENDIX F. DUPLICATE FIGURES IN METRIC UNITS

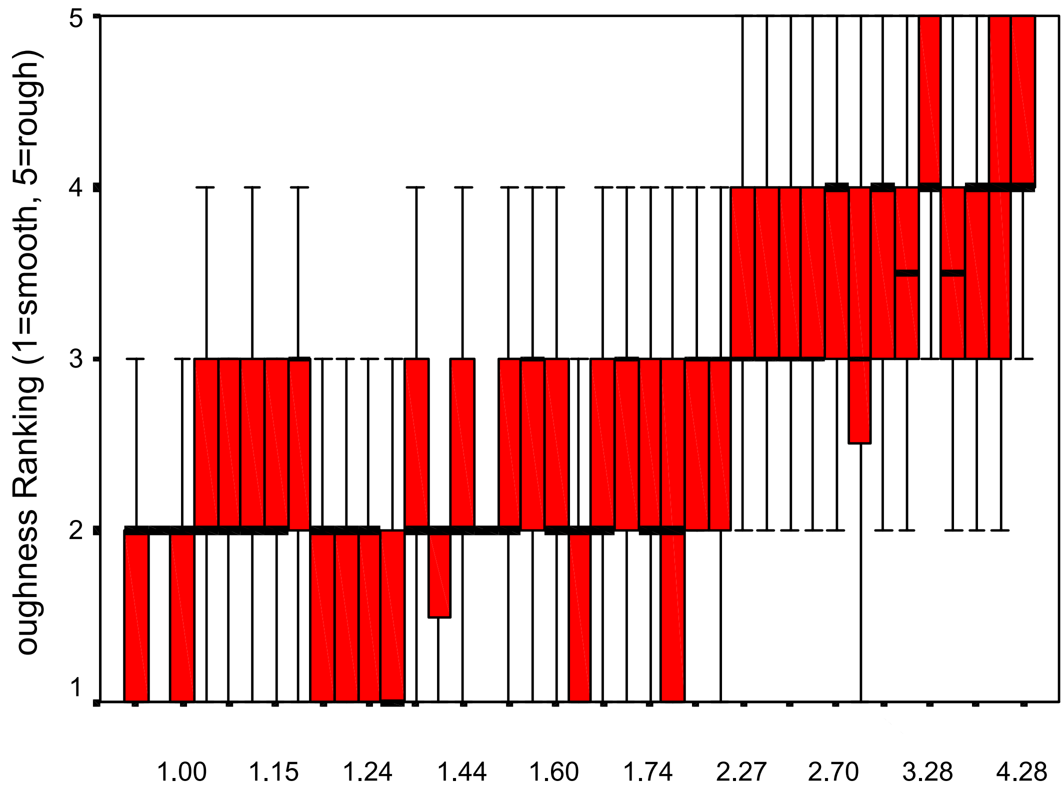


Figure F-1. Boxplot of Roughness Ranking vs. Measured IRI (m/km)

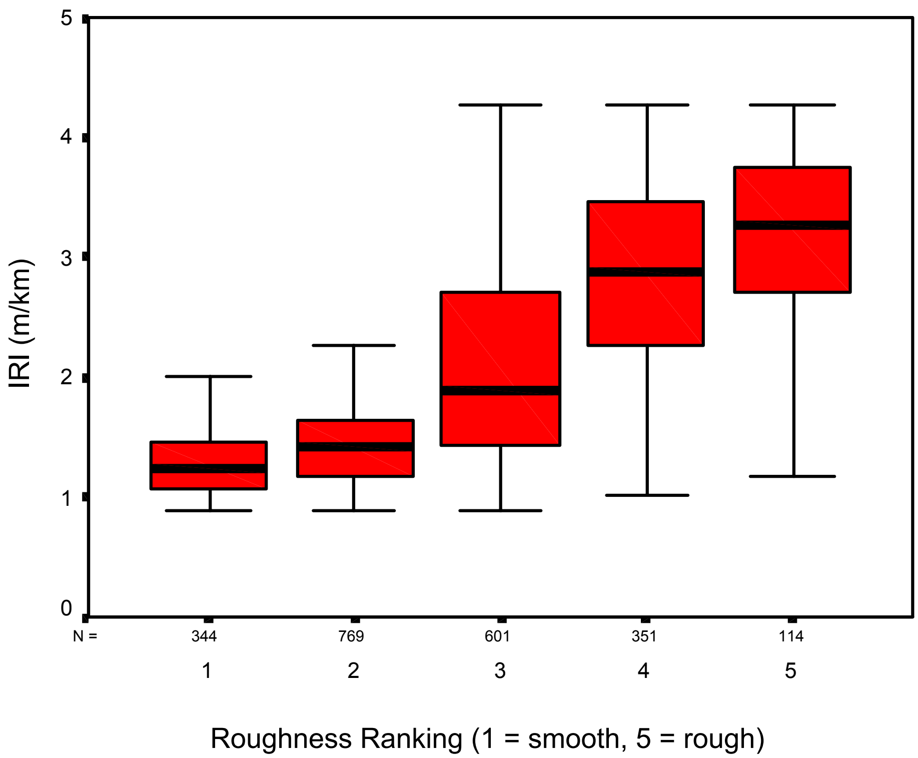


Figure F-2. Boxplot of IRI (m/km) vs. Roughness Ranking

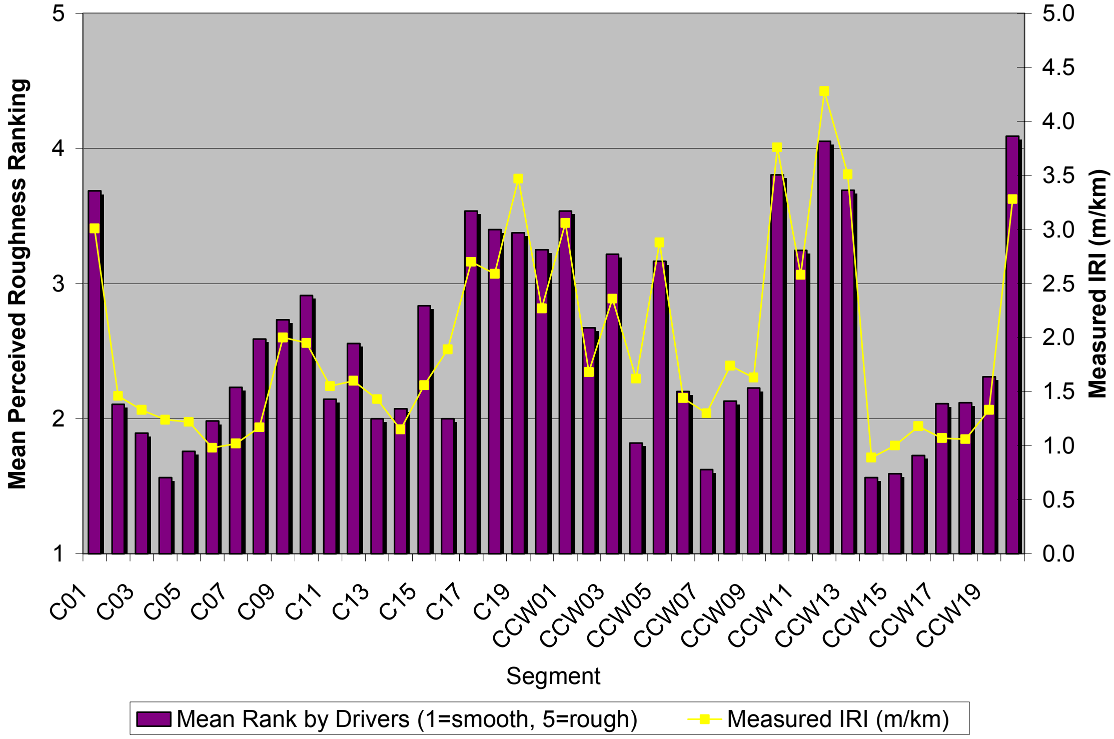


Figure F-3. IRI Measurements (m/km) vs. Driver Roughness Rankings

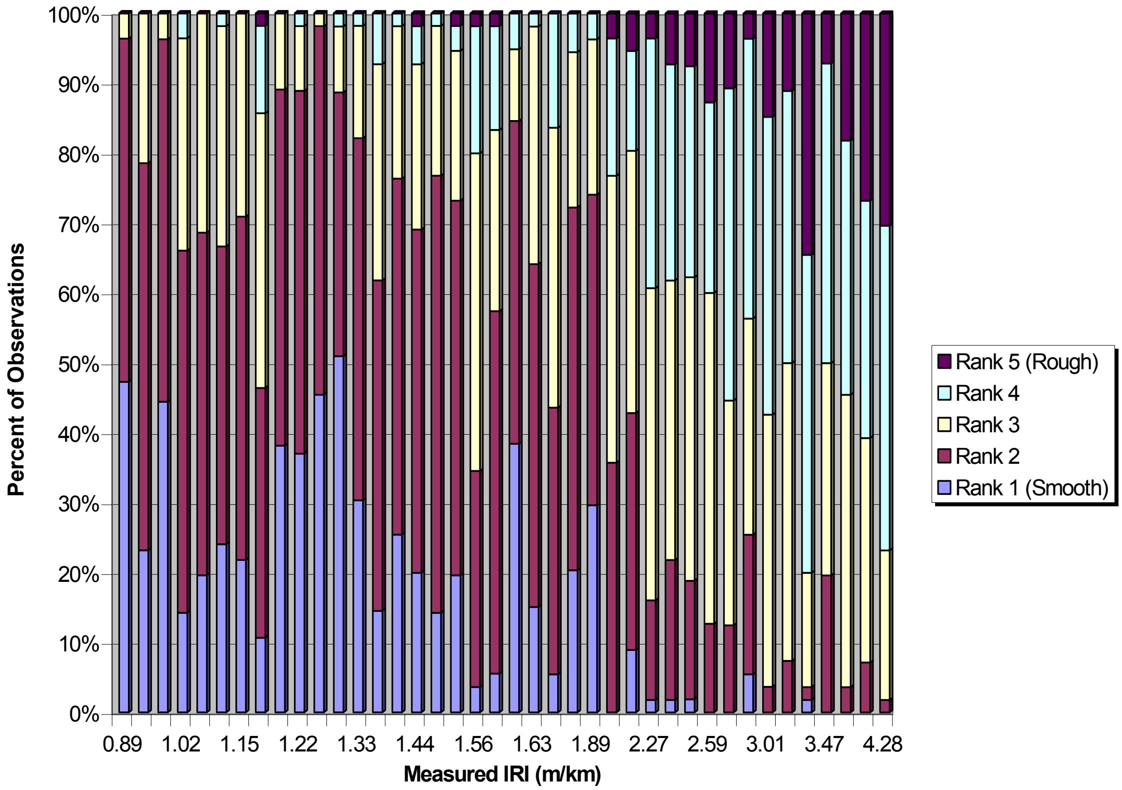


Figure F-4. Cumulative Percentage Plot of Roughness Rankings on Each Test Section vs. Corresponding Ordered IRI Measurements (m/km)

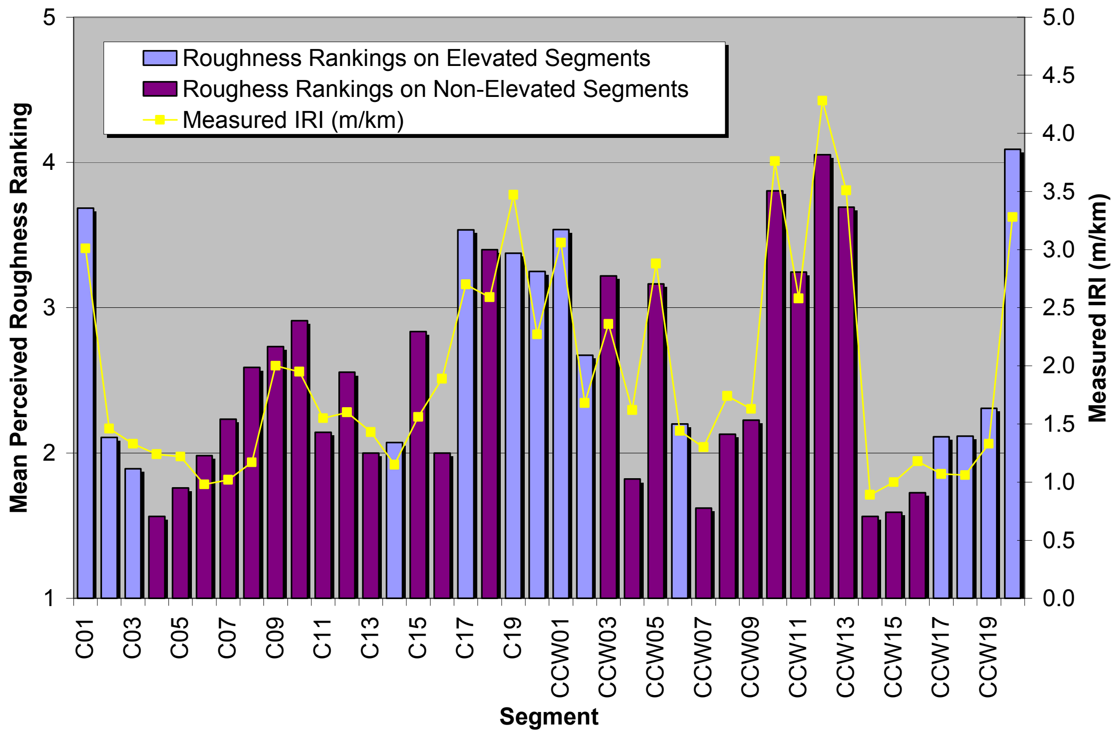


Figure F-5. IRI Measurements (m/km) and Average Roughness Rankings on Elevated Segments

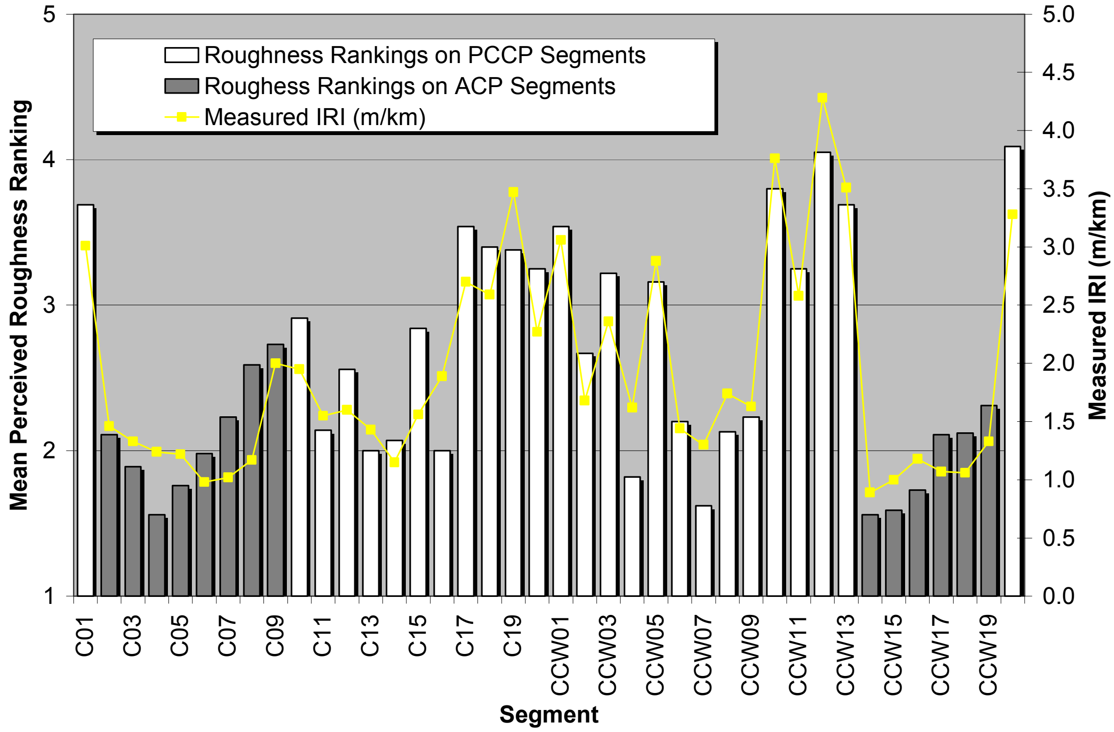


Figure F-6. IRI Measurements (m/km) and Average Roughness Rankings by Surface Type

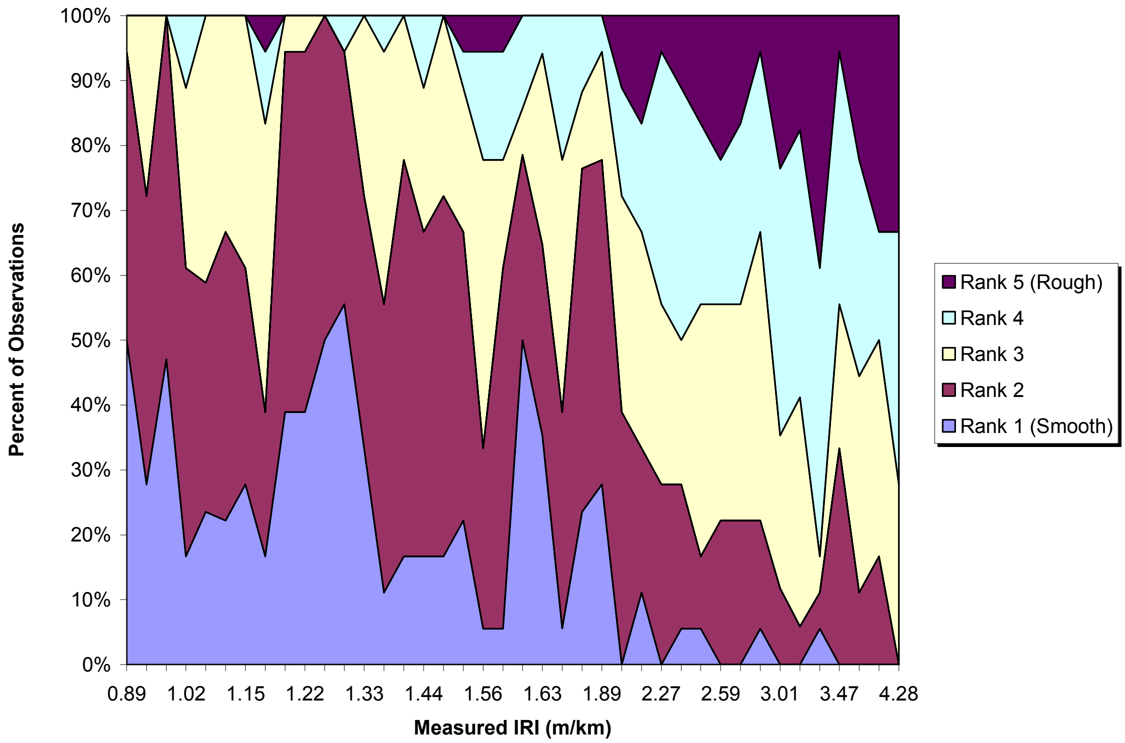


Figure F-7. Roughness Ranking vs. Measured IRI (m/km) for Subset of Drivers in Midsize Sedan Test Vehicle

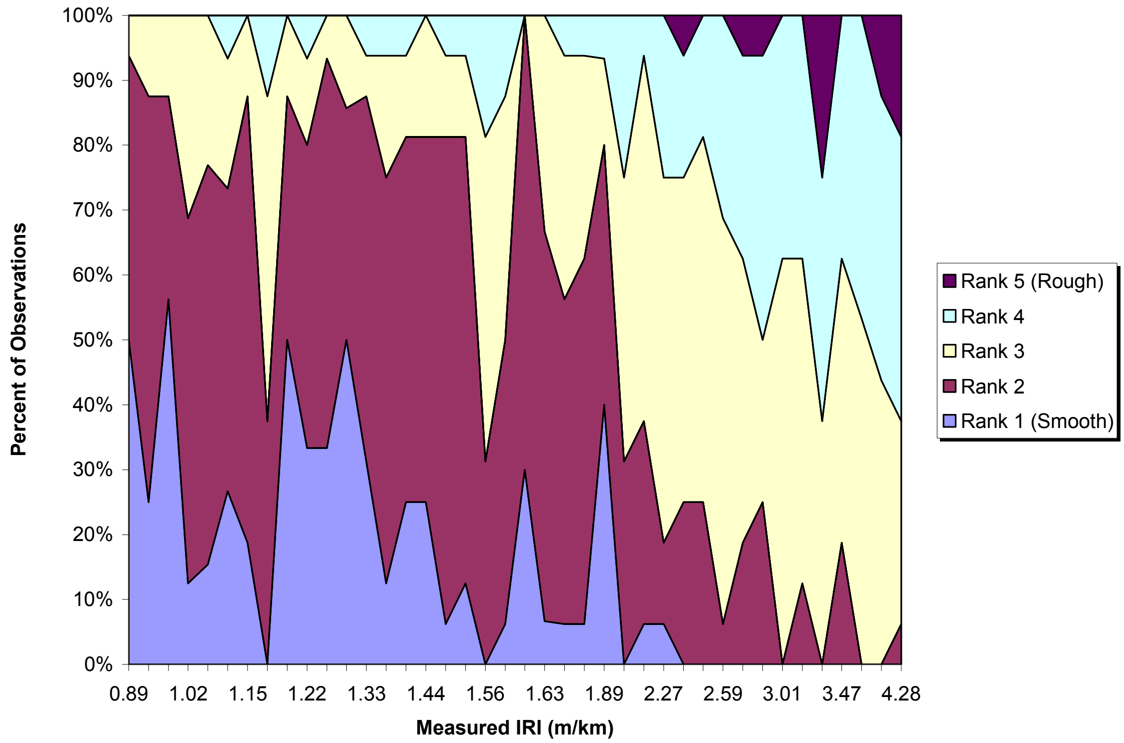


Figure F-8. Roughness Ranking vs. Measured IRI (m/km) for Subset of Drivers in Sports Utility Test Vehicle

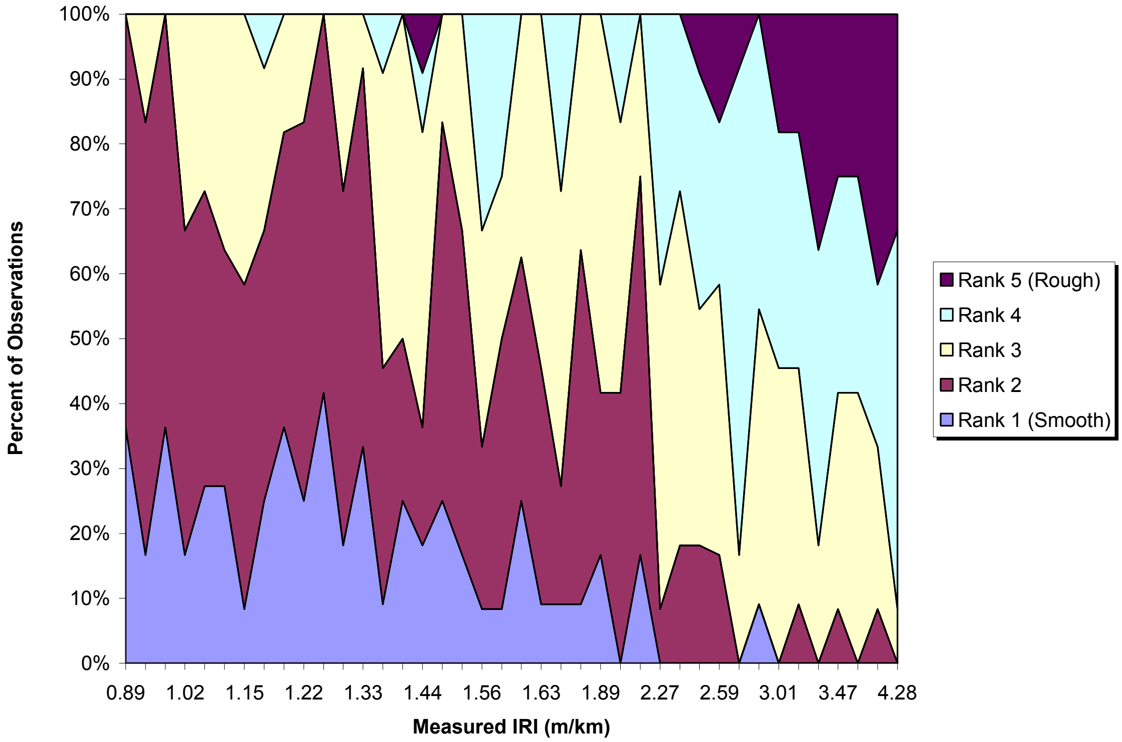


Figure F-9. Roughness Ranking vs. Measured IRI (m/km) for Subset of Drivers in Pickup Truck Test Vehicle

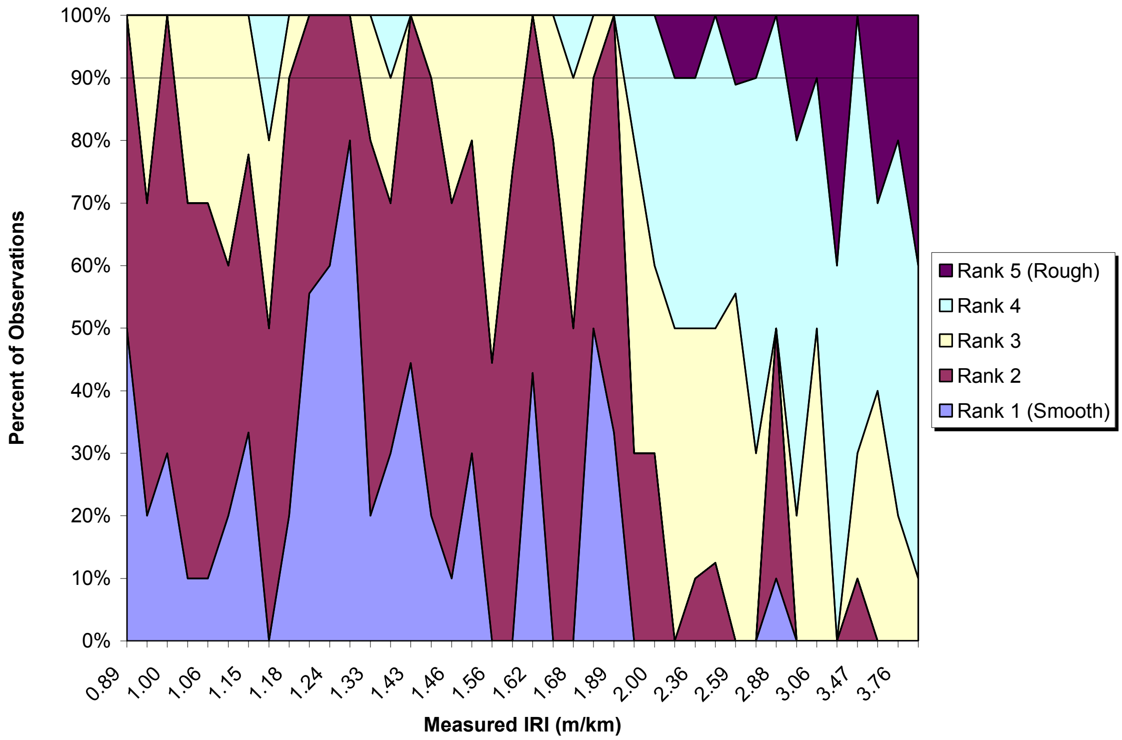


Figure F-10. Roughness Ranking vs. Measured IRI (m/km) for Subset of Drivers in Minivan Test Vehicle

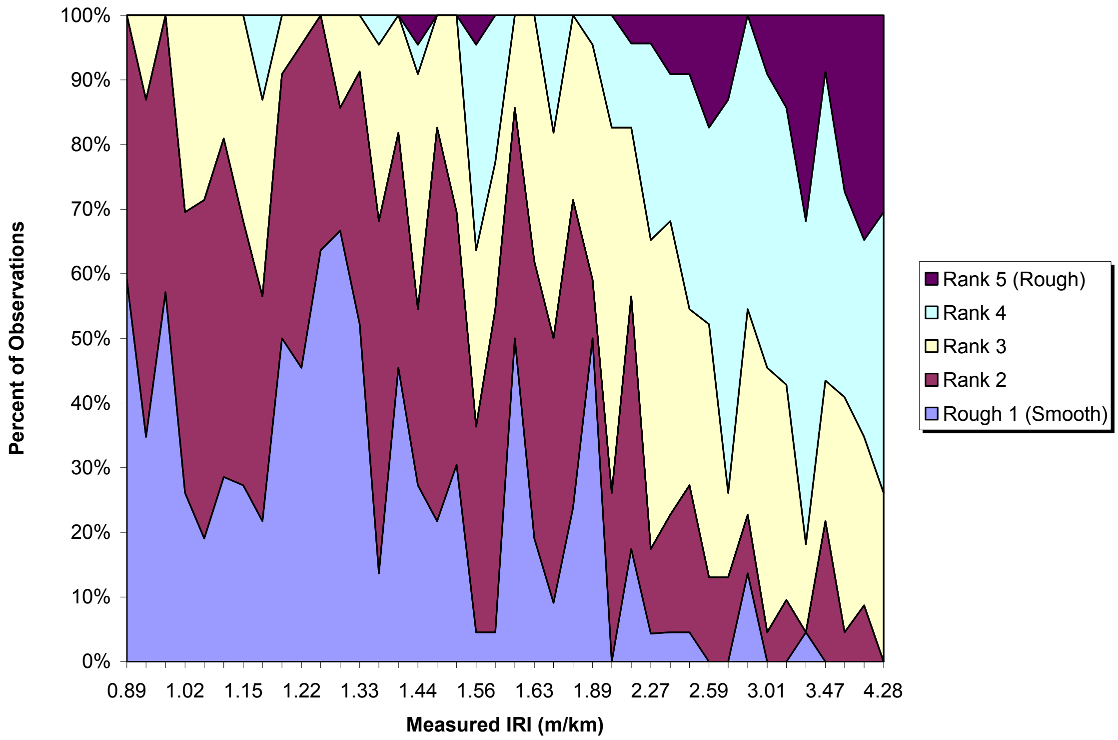


Figure F-11. Roughness Ranking vs. Measured IRI for Subset of Female Drivers

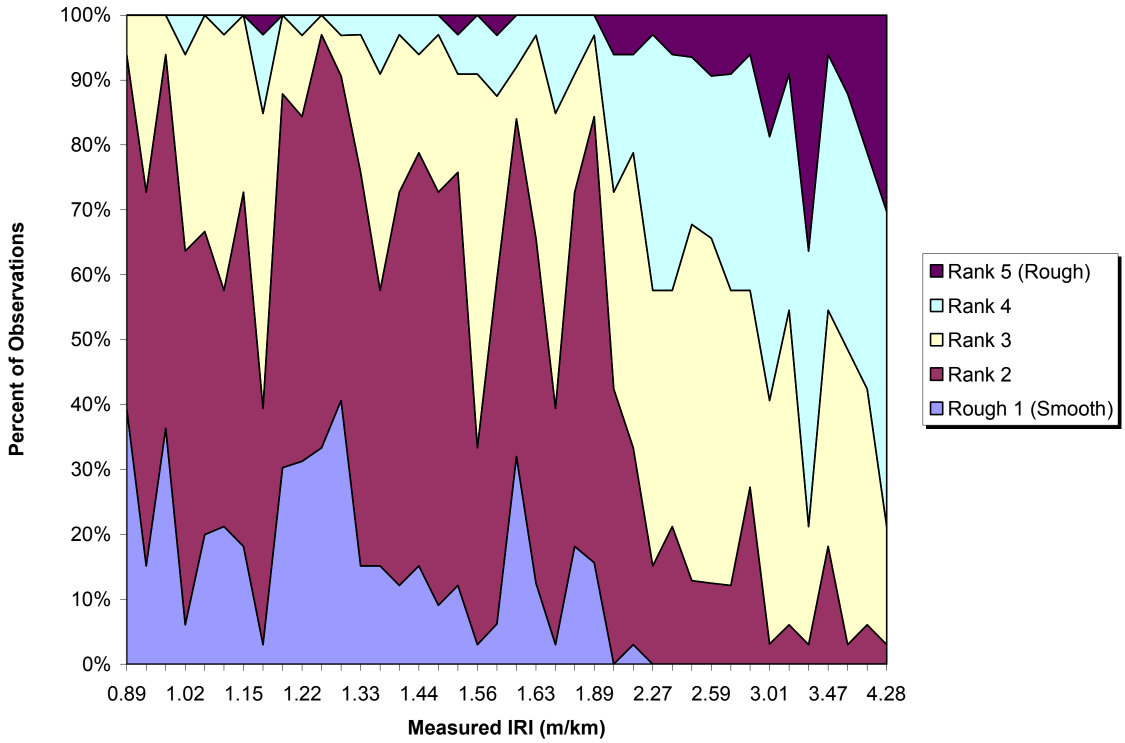


Figure F-12. Roughness Ranking vs. Measured IRI (m/km) for Subset of Male Drivers